PCT/PTO 21 JUL 2000

FORM PTO-1390 (REV 10-98) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C.371		**UMV-1474US**
		U.S. APPLICATION NO. (If known, see 37 CFR 1.5) **09/600766**
INTERNATIONAL APPLICATION **PCT/US99/01382**	INTERNATIONAL FILING DATE **21 January 1999 (21.01.99)**	PRIORITY DATE CLAIMED **21 January 1998 (21.01.98)**

TITLE OF INVENTION

TARGETING GENE TRANSFER VECTORS TO CERTAIN CELL TYPES BY PSEUDOTYPING WITH VIRAL GLYCOPROTEIN

APPLICANT(S) FOR DO/EO/US

Gary J. NABEL; Rafael DELGADO; and Zhi-yong YANG

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C.371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This express request to begin national examination procedures (35 U.S.C. 371 (f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371 (b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☒ has been transmitted by the International Bureau.
   c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   b. ☐ have been transmitted by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)) **(unexecuted) (4 sheets);**
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included
13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information: **Transmittal Letter (2 sheets in duplicate); Response to the Invitation to Correct Defects in the International Application with Sequence Listing attached along with Sequence listing diskette (34 sheets); PCT International Published Application (WO 99/37331) (with International Search Report); (76 sheets); Check ($840); Certificate of Express Mailing (1 sheet); and Return Postcard.**

Rec'd PCT/PTO 21 JUL 2000

U.S. APPLICATION NO. (if known, see 37 CFR 1.5)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NO.
09/600766	PCT/US99/01382	UMV-1474US

17. ☒ The following fees are submitted:

	CALCULATIONS PTO USE ONLY	
BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) .(a/o January 1, 2000):		
Search Report has been prepared by the EPO or JPO.....$970		
International preliminary examination fee paid to USPTO (37 CFR 1.482).....$840		
No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)).....$690		
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO.....$670		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4).....$96		
ENTER APPROPRIATE BASIC FEE AMOUNT =	**$840**	
Surcharge of **$130.00** for furnishing the oath or declaration later than 20 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$--	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	**20** -20 =	**0**	X **$18.00**	**$0**	
Independent claims	**2** -3 =	**0**	X **$78.00**	**$0**	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **260.00**	$--	
TOTAL OF ABOVE CALCULATIONS =				**$840**	
Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28)				$--	
SUBTOTAL =				**$840**	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f)). +				$--	
TOTAL NATIONAL FEE =				**$840**	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$--	
TOTAL FEES ENCLOSED =				**$840**	
				Amount to be: refunded	$
				charged	$

a. ☒ A check in the amount of **$840** to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. _____ in the amount of $_____ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. **12-0080**. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

DeAnn F. Smith
LAHIVE & COCKFIELD, LLP
28 State Street
Boston, Massachusetts 02109
United States of America
(617)227-7400
Date: 21 July 2000

SIGNATURE

DeAnn F. Smith
NAME

36,683
REGISTRATION NUMBER

534 Rec'd PCT/PTC 21 JUL 2000

TARGETING GENE TRANSFER VECTORS TO CERTAIN CELL TYPES BY PSEUDOTYPING WITH VIRAL GLYCOPROTEIN

FIELD OF THE INVENTION

The present invention relates generally to compositions and methods for selective gene transfer, and in particular, to methods for targeting genes to certain cell types, comprising introducing to a cell population the gene to be transferred operatively-linked to an appropriate transfer vehicle, wherein the transfer vehicle is associated with a transmembrane form of viral glycoprotein.

BACKGROUND OF THE INVENTION

Ebola virus has been identified as the cause of several highly lethal outbreaks of hemorrhagic fever. Infection begins typically with flu-like symptoms which often progress rapidly to fatal complications of hemorrhage, fever, and hypotensive shock. Bowen, E.T.W. et al., *Lancet* 1:571 (1977); Centers for Disease Control, *M.M.W.R.* 44:381 (1995); Le Guenno, B. et al., *Lancet* 345:1271 (1995); Peters, C.J. et al., *Fields Virology*, B.N. Fields, D.M. Knipe and P.M. Howley, Eds. (Lippincott-Raven, Philadelphia) p. 1161 (1996). The negative-stranded genome of Ebola virus contains seven structural and regulatory proteins (Sanchez, A. et al., *Virus Res.* 29:215 (1993)), but despite its relative simplicity, the molecular basis for Ebola virus pathogenicity is unknown. Among the viral gene products, the glycoprotein is found in two forms: a secreted form, 50-70 kD (Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996)), synthesized at high levels early in the course of infection, and an alternative transmembrane form, which arises from RNA editing to encode a 120-150 kD glycoprotein that is incorporated into the virion. Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996); Volchkov, V.E. et al., *Virology* 214:421 (1995). The first 295 amino acids (aa) of both proteins are identical in the Zaire strain, while sGP contains an additional 69 and GP another 381 COOH- terminal aa residues. Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996). The specific cellular targets of these related gene products and their roles in the pathogenesis of Ebola infection have not been characterized.

SUMMARY OF THE INVENTION

The present invention provides compositions and methods for targeting gene transfer vectors to certain cell types by pseudotyping with a transmembrane form of viral glycoprotein. In one embodiment, the methods of the invention comprise the

step of administering to a cell population a gene to be transferred operatively-linked to an appropriate transfer vehicle, wherein the transfer vehicle is associated with a transmembrane form of Ebola glycoprotein. In this embodiment, the gene will be targeted to cell types naturally infected with Ebola such as endothelial cells, monocytes and hepatocytes.

Genetic constructs for selective gene transfer into certain cell types are also provided. The genetic constructs of the present invention comprise a gene to be transferred operatively-linked to an appropriate transfer vehicle or carrier, wherein the transfe vehicle or carrier is associated with a transmembrane form of viral glycoprotein. In one embodiment, the transmembrane form of Ebola glycoprotein is expressed on the surface of a virus-based gene-targeting vector, *e.g.*, lentiviral or retroviral vector. In another embodiment, an expressed or synthesized transmembrane glycoprotein is chemically derivatized to a non-biologic gene targeting vehicle.

Additional objects, advantages, and features of the present invention will become apparent from the following description and appended claims, taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The various advantages of the present invention will become apparent to one skilled in the art by reading the following specification and subjoined claims and by referencing the following drawings.

Figures 1A-1C show the binding of sGP to neutrophils;

Figures 2A-2D show the infection of different cell types by a GP-pseudotyped vector of the present invention;

Figures 3A-3F show the dependence of sGP binding on CD16b and correlation of binding with neutrophil activation;

Figures 4A-4B show the effect of sGP on neutrophil function;

Figures 5A-5C show the infection rate of cells with a GP-pseudotyped retroviral vector of the present invention;

Figure 6 is a schematic of the plasmid pVR 1012-GP(IC) (Ivory Coast strain of GP, see SEQ ID NO: 1);

Figure 7 is a schematic of the plasmid pVR 1012-GP(S) (Sudan strain of GP, see SEQ ID NO: 2);

Figure 8 is a schematic of the plasmid pVR 1012-GP(Z) (Zaire strain of GP, see SEQ ID NO: 3);

Figure 9 is a schematic of the plasmid pVR 1012-sGP(Z) (Zaire strain of sGP, see SEQ ID NO: 4); and

Figure 10 is a summary of the characterization of GP and sGP derivatives for their ability to pseudotype to induce cytotoxicity in producer cells.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention provides genetic constructs and methods for targeting gene transfer vectors to certain cell types by pseudotyping with a transmembrane form of viral glycoprotein. The methods for selective gene transfer of the present invention comprise the step of administering to a cell population a genetic construct of the present invention so that the gene is transferred and expressed in certain cell types present in the cell population. Administration to the cell population may be *ex vivo* or *in vivo*.

The genetic constructs of the present invention comprise a gene to be transferred operatively-linked to an appropriate transfer vehicle or carrier, wherein the transfer vehicle or carrier is associated with a transmembrane form of viral glycoprotein. In one embodiment, the transmembrane form of Ebola glycoprotein is associated with the vehicle or carrier. The gene to be transferred will thus be targeted to cell types naturally infected with Ebola virus including endothelial cells, hepatocytes, monocytes and related cell types such as dendritic cells. The transmembrane form of Ebola glycoprotein may be chosen from, without limitation, the Ivory Coast strain (SEQ ID NO: 1), Sudan strain (SEQ ID NO: 2), the Zaire strain (SEQ ID NO: 3) and/or the Reston strain. It will be appreciated that in other embodiments of the present invention, other hemorrhagic fever virus glycoproteins, in particular transmembrane glycoproteins, may be employed and will target those cell types naturally infected by the virus. Examples of hemorrhagic viruses include dengue virus, Yellow Fever virus (*flaviviridae*); Lassa, Junin and Machupo (*arenaviridae*); Rift Valley, Congo-Crimean and Hantaan (*bunyaviridae*); and Marburg (*filoviridae*). It will also be appreciated that derivatives of the transmembrane glycoprotein which retain the capability of targeting specific cell types, may also be employed, for example, the transmembrane glycoproteins may be mutated, *e.g.*, toxic regions may be removed to improve producer cell viability (see Figure 10).

The transmembrane glycoprotein may be expressed on the surface of a virus-based gene-targeting vector, *e.g.*, lentiviral, retroviral, replication-deficient retroviral, adenoviral or adeno-associated viral vector. The transmembrane glycoprotein may

also be expressed or synthesized and chemically derivatized to a non-biologic gene targeting vehicle, *e.g.*, liposome or DNA-protein complex.

The term "operatively-linked" as used herein refers to functional linkage between a nucleic acid expression control sequence (such as a promoter) and a second nucleic acid sequence (*i.e.*, gene), wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence. Expression control sequences are known to those skilled in the art (see, *e.g.*, Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990)). "Associated with" as used herein refers to the transmembrane form of viral glycoprotein being in contact or linkage with the transfer vehicle or carrier in such a way as to direct the transfer vehicle or carrier to certain cell types. The terms "transfer vehicle" and "carrier" refer to any type of structure which is capable of delivering the gene of interest to a target cell.

Many transfer vehicles or carriers are known in the art. For example, various viruses that are capable of infecting cells can be recombinantly manipulated to carry the gene of interest without affecting their infectivity. As used herein, the terms "infect" and "infectivity" refer only to the ability of a virus to transfer genetic material to a target cell. Those terms do not mean that the virus is capable of replication in the target cell. In fact, it is preferable that such viruses are replication defective so that target cells do not suffer the effects of viral replication.

In one embodiment, the virus employed is a replication defective retroviruses. When these replication defective retroviruses are employed, their genomes can be packaged by a helper virus in accordance with well-known techniques. Suitable retroviruses include PLJ, pZip, pWe and pEM, each of which is well known in the art. Suitable helper viruses for packaging genomes include ψCrip, ψCre, ψ2, ψAm and adeno-associated viruses.

In another embodiment, lentiviral vectors are employed. Surprisingly, the inventors of the present invention were successful in pseudotyping lentiviral vectors (HIV) with the transmembrane glycoprotein from Ebola. Feline immunodeficiency virus, bovine immunodeficiency virus, simian immunodeficiency virus and EAIV, may also be employed as the carrier in the compositions and methods of the present invention.

Gene delivery systems other than viruses may also be employed. For example, the gene to be transferred may be packaged in a liposome which is chemically derivatized to the transmembrane glycoprotein. To form these liposomes,

one mixes the DNA of an expression vector which expresses the gene to be transferred with lipid, such as *N*-[1-(2,3,-dioleyloxy)propyl]-*N,N,N*-trimethylammonium chloride (DOTMA) in a suitable buffer, such as Hepes buffered saline. This causes the spontaneous formation of lipid-DNA complexes (liposomes). Felgner, P.L. et al., *PNAS (USA)* 84:7413-7417 (1987).

Another gene delivery system that may be utilized in this invention is DNA-protein complexes. The formation of DNA-protein complexes is described in United States Patent No. 5,166,320, the disclosure of which is herein incorporated by reference.

It will be appreciated that any gene may be employed in the compositions and methods of the present invention. For example, and without limitation, in the treatment of cancer, death inducing genes, including genes coding for cytostatic or cytotoxic proteins, *e.g.*, HSV tk, and genes encoding cyclin dependent kinase inhibitors, p21, p27, cytosine deaminase, and fas ligand, may all be employed. In another example, for the treatment of cardiovascular or ischemic vascular disease, genes encoding angiogenic factors such as VEGF basic or acidic FGF's (FGF 1-5) may be employed. In yet another example, in the treatment of vasospasm, the gene encoding NO synthase or heme oxygenase, may be employed. In a further example, monocytes and dendritic cells may be targeted with genes encoding immunogens for cell-targeted immunization.

In one embodiment, the methods of targeting gene transfer vectors to certain cell types involve administering to a cell population *ex vivo*, a construct of the present invention and introducing the transfected cells into a subject. In an alternative embodiment, the methods of the present invention comprise administering to an *in vivo* cell population a construct of the present invention. Administration can be by any of the routes normally used for *in vivo* gene therapy such as direct delivery to cells via a gene gun, and other known techniques. The constructs are thus administered in any suitable manner, preferably with pharmaceutically acceptable carriers. The constructs can be administered, for example, by intravenous infusion, orally, topically, intraperitoneally, intravesically or intrathecally. The preferred method of administration will often be intravenous.

To practice an *ex vivo* method of the present invention, a source of cells is obtained. The cells are optionally selected from *in vitro* cells, such as those derived from cell culture and *ex vivo* cells, such as those derived from a subject. The term "subject" is intended to include living organisms, *e.g.*, mammals. Examples of

subjects include humans, primates, dogs, cats, mice, rats, and transgenic species thereof. It will be appreciated that specific cell populations may be obtained by isolation from certain tissues by methods known to those skilled in the art. The cells are maintained under conditions necessary to support growth, for example an appropriate temperature (*e.g.*, 37°C) and atmosphere (*e.g.*, air plus 5% CO_2).

The cells are then transfected with the constructs of the present invention by introducing the constructs to the cell population, under conditions favorable for transfection. According to one embodiment of the present invention, cells are treated with compounds that facilitate uptake of the constructs by the cells. According to another embodiment of the present invention, cells are treated with compounds that stimulate cell division and facilitate uptake of the constructs. It will be appreciated that compounds that facilitate uptake of constructs by cells and compounds that stimulate cell division are known to those skilled in the art.

The constructs of the present invention express the transferred gene in a dose dependent manner. The specific dose to be administered to a patient will be determined by the efficacy of the particular construct and/or delivery system employed, the gene transferred, and the condition of the patient, as well as the body weight or surface area of the patient to be treated. The size of the dose also will be determined by the existence, nature, and extent of any adverse side-effects that accompany the administration of a particular construct or effect a particular patient. In determining the effective amount of the construct or transfected cell to be administered, the physician needs to evaluate circulating plasma levels, toxicities, and progression of disease. It will be appreciated that administration can be accomplished via single or divided doses.

There is a wide variety of suitable formulations for pharmaceutical compositions containing the constructs of the present invention. Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the construct dissolved in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the construct, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. The construct, alone or in combination with other suitable components, may also be made into aerosol formulations to be administered via inhalation, *e.g.*, to the bronchial passageways. Aerosol formulations can be placed into pressurized acceptable propellants, such as dichlorodifluoromethane, propane, nitrogen, and the like.

Suitable formulations for rectal administration include, for example, suppositories, which consist of the construct with a suppository base. Suitable suppository bases include natural or synthetic triglycerides or paraffin hydrocarbons. In addition, it is also possible to use gelatin rectal capsules which consist of a combination of the construct with a base, including, for example, liquid triglycerides, polyethylene glycols, and paraffin hydrocarbons.

Formulations suitable for parenteral administration, such as, for example, by intraarticular (in the joints), intravenous, intramuscular, intradermal, intraperitoneal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. The formulations can be presented in unit-dose or multi-dose sealed containers, such as ampules or vials. Extemporaneous injection solutions and suspensions can be prepared from sterile powders, granules, and tablets of the kind previously described. Cells transfected by the constructs as described above in the context of *ex vivo* therapy can also be administered as described above.

This invention also provides compositions and kits comprising the constructs of the present invention. For example, the composition can comprise the constructs of the present invention in a pharmaceutically acceptable carrier as described above. Kits comprising such compositions and instructions for use are also within the scope of this invention.

In order to more fully demonstrate the advantages arising from the present invention, the following examples are set forth. It is to be understood that the following is by way of example only and is not intended as a limitation on the scope of the invention.

SPECIFIC EXAMPLE 1

I. Methods

Recombinant retroviruses were produced by transient transfection of 293T cells: 2 x 10^6 cells were plated 24 hours before transfection in 60 mm dishes. Transfection was performed by calcium-phosphate precipitation using 3 μg of a retroviral vector (Kinsella, T.M. et al., *Hum. Gene Ther.* 7:1405 (1996)) encoding lucerifase linked to an internal ribosome entry site and a green fluorescent protein derivative (GFP; pEGFP, Clontech), pLZRs-Luc-Gfp, 5 μg of an expression vector

encoding gag and pol, pNGVL-MLVgag-pol, and 1 μg of the envelope encoding plasmid: pNGVL-4070A (ampho) env, pCMV-Eco env or p1012-Ebola GP, respectively. Supernatants corresponding to 24-48 hours post-transfection were harvested, cleared by low-speed centrifugation and either used immediately for infection or frozen at -80°C. Infections were performed in 6-well plates (1.5-2.5 x 10^5 adherent cells) or 12-well plates (5 x 10^5 non-adherent) using different dilutions of the supernatants by incubating the cells overnight with 1 ml and 300 μl, respectively of the diluted supernatants. Polybrene was used at a concentration of 5 μg/ml for all the cell lines except for D17 in which the concentration was 100 μg/ml. After overnight infection, fresh medium was added and the cells were incubated for an additional 24 hours. After infection, the cells were lysed in 25 mM Tris-phosphate pH 8, 2 mM DTT, 2 mM 1,2-diaminocyclohene-N,N,N',N'-tetraacetic acid, 10% glycerol, 1% TritonX-100, and assayed for luciferase activity using Luciferase Assay Reagent (Promega, Madison, WI) in a 1251 BioOrbit Luminometer. The same number of cells (range 5-10 x 10^4) was analyzed for every specific cell line.

Binding of sGP to neutrophils and inverse correlation of binding with activation: *Figures 1A-1A2.* PBMC from normal volunteers were incubated with control or sGP supernatants derived from transfected 293 cells, and immunostaining was performed using a rabbit antibody to sGP as previously described. Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996); Xu, L. et al., *Nat. Med.* (1997) in press. Secondary staining was performed with a fluorescein isothiocyanate (FITC)-conjugated goat anti-rabbit IgG antibody (Sigma, F9887). All incubations were performed at 4°C for 30 minutes with .4 μg of the relevant antibodies per 10^6 cells in a 50 μl volume.

Figures 1B-1B1. Double immunostaining with antibodies to sGP and the neutrophil-specific marker, CD15. Cells were incubated with a FITC conjugated mouse anti-human CD15 antibody (Caltag, cat# MHCD1501), followed by secondary staining with a PE-conjugated anti-rabbit IgG antibody (Sigma) to detect sGP binding. Cells were washed with PBS, fixed in 1% formaldehyde, and analyzed by FACS.

Figure 1C. Specific absorption of sGP by neutrophils. Control or sGP supernatants derived from relevant transfected 293 cells (Xu, L. et al., *Nat. Med.* (1997) in press) were incubated at 1:500 dilution with 10^6 mononuclear or granulocytic cells. Cells were removed and the resulting supernatants analyzed by an 8% SDS PAGE gel. Western blot analysis was performed as previously described (Xu, L. et al., *Nat. Med.* (1997) in press) using an anti-GP rabbit antisera and a secondary

antibody, horseradish peroxidase conjugated donkey anti-rabbit IgG at a dilution of 1:5,000 (Amersham, NA934). Primary antibody was incubated for 30 minutes at room temperature, as was the secondary antibody. The immunocomplexes were detected by chemiluminescence using Supersignal® chemiluminescent substrate reagents (Pierce) according to the manufacturer's instructions. Arrow indicates sGP reactive band.

Infection of different cell types by GP-pseudotyped retroviral vector and preferential binding to endothelial cells: *Figure 2A.* Infection of different indicator cell lines with the Ebola-GP pseudotyped retrovirus expressing luciferase. Amphotropic and ecotropic retroviral vectors were used as controls. Viruses were diluted to different multiplicities of infection (MOI) to provide for equal luciferase activity on Hela cervical epithelial cells, permissive for amphotropic retrovirus, D17 dog osteosarcoma cells (Embretson, J.E. et al., *J. Virol.* 61:3454 (1987)), which are permissive for amphotropic, xenotropic, and ecotropic retroviruses, and BW5147 T leukemia cells permissive for amphotropic and ecotropic virus. In these groups, GP virus titer was 1-4x10^5/ml and amphotropic virus was ~2x10^4/ml (MOI's $\approx$ 1.0 and 0.1, respectively), and the ecotropic virus titer was ~10^6/ml (MOI $\approx$ 10). Titers were determined by endpoint dilution of reporter activity of the amphotropic virus in D17 cells, and was normalized to reverse transcriptase activity for the GP virus.

Figure 2B. Analysis of different normal or transformed cell lines by infection with amphotropic or GP retroviral vectors at the same titer (10^4/ml, MOI $\approx$ 0.2). Forty-eight hours after infection, an equivalent of 5 x 10^4 cells was assayed for luciferase activity after exposure to equal titers of viral stocks. Luminescence is expressed as the fold-increase over non-infected control cells.

Figures 2C-2C3. The binding of sGP (left) or GP-pseudotyped retrovirus (right) to neutrophils (upper panel) or microvascular endothelium (lower panel) was determined by FACS. sGP binding was performed as in Fig. 1A, and retrovirus incubation was performed at 37°C for 2 hours in the presence of polybrene (8 μg/ml).

Figure 2D. Infection of D17 cells by GP-pseudotyped virus in the absence (lane 1, none) or presence of control (lane 2) or sGP supernatant (lane 3) from transfected 293 cells. Gene transfer was measured by the luciferase assay as described below. Luminescence refers to relative light units in the luciferase assay.

Dependence of sGP binding on CD16b and correlation of binding with neutrophil activation: *Figures 3A-3D.* Neutrophils were incubated for 30 minutes at 4°C with a mouse antibody to CD16b (upper panel; clone 3G8 from Immunotech,

cat# 1M0813) or CD62L (middle panel, R&D Systems), compared to the indicated control antibody [purified mouse IgG (Vector Laboratories), I-2000], followed by supernatants from control or sGP-transfected 293 cells, primary rabbit antibody to sGP, and a FITC-conjugated secondary antibody to rabbit IgG (Fig. 1, legend). Cells were washed with PBS, fixed in 1% formaldehyde, and analyzed by FACS. For blocking, 10^6 cells were incubated with 0.5 – 1 μg of the relevant antibodies for 30 minutes in a 50 μl volume.

Figures 3E-3F. Immunostaining with sGP was performed on isolated neutrophils which were maintained in media (none) or incubated with PMA (10 ng/ml) at 37°C for 30 minutes (PMA).

Effect of sGP on neutrophil function: *Figures 4A-4B.* Exposure of neutrophils to sGP inhibits down modulation of L-selectin. Isolated neutrophils were incubated with the indicated control or sGP containing supernatants (Xu, L. et al., *Nat. Med.* (1997) in press) and defined media (AIM V, GIBCO) for 4 hours at 37°C. Expression of L-selectin was determined using an anti-CD62L antibody (R&D Systems), followed by the secondary staining using a FITC-conjugated anti-mouse IgG (Sigma, F2883) as described in Fig. 1, legend. Cells were washed with PBS, fixed with 1% formaldehyde and analyzed by FACS for relative levels of fluorescence intensity as a function of cell number. An isotype control was used to quantitate background levels of immunostaining (neg.). Results are representative of three independent experiments.

II. Results

To determine the specificity of Ebola virus glycoproteins, expression vectors encoding either sGP, GP, or a plasmid control (Xu, L. et al., *Nat. Med.* (1997) in press) were transfected into 293 cells, and cell culture supernatants were used as a source of relevant recombinant glycoproteins. Binding of sGP was determined by immunofluorescence analysis after incubation of relevant supernatants with normal or transformed human cell lines. No binding was detected to several hematopoietic lineages, including lymphocytes or monocytes (Fig. 1A), or transformed Jurkat or CEM T leukemias, the HL60 myelomonocytic or U937 promonocytic leukemia cells. In contrast, sGP was able to bind to granulocytic cells, as evidenced by FACS analysis of this subset of peripheral blood mononuclear cells (PBMC) discriminated by cell size and granularity (Fig. 1A). This cell specificity was confirmed by using double-staining with a granulocyte-specific cell surface marker, CD15 (Fig. 1B). Absorption of sGP

by purified neutrophils in the absence of antibodies also resulted in depletion of sGP, indicating that binding to the neutrophil occurred in the absence of antibody (Fig. 1C).

A potential structural similarity between Ebola GP and avian sarcoma virus envelope protein has been previously proposed (Gallaher, W.R., *Cell* 85:477 (1996)), raising the possibility that this protein could be incorporated into retroviral particles. To determine the binding specificity of the transmembrane glycoprotein, pseudotyping of a Moloney leukemia virus was therefore attempted. Infectivity of different cell types by this pseudotyped vector was determined with a luciferase reporter gene. de Wet, J.R. et al., *Mol. Cell. Biol.* 7:725 (1987). This analysis revealed infection of cells different from those which interacted with sGP (Fig. 2A,B). For example, though it could infect other cell types, transduction by the GP retroviral vector readily occurred in endothelial cells, either from the microvasculature (MVEC) or umbilical veins (HUVEC) (Fig. 2B), which did not bind sGP (Fig. 2C, left). When the specificity of GP-retrovirus was compared to murine retroviruses pseudotyped with amphotropic or ecotropic envelope gp70 proteins, the range of susceptible target cells differed (Fig. 2B), suggesting that the virus receptor(s) for Ebola GP differ from those previously described for gp70. Minimal binding of GP-virus was observed on neutrophils, despite the ability of these cells to bind sGP (Fig. 2C, upper panel) and the fact that immunoreactive protein was detected on the virus. Conversely, GP-virus binding to endothelial cells was readily detected, though these cells did not bind sGP (Fig. 2C, lower panel). When sGP was analyzed for its effect on GP retroviral gene transfer, infection was not inhibited by sGP (Fig. 2D), further confirming the divergent specificities of the two forms of the viral glycoprotein. Recent studies have revealed that the biochemical forms of these proteins differ, with sGP present in solution primarily as a homodimer and GP as a trimer, suggesting that differences in multimer composition may contribute to these alternative specificities.

Potential cell surface receptors for sGP were analyzed with antibodies to several neutrophil cell surface antigens to interfere with sGP binding, including CD15, L-selectin (CD62L), CD16b, and several common leukocyte antigens. Only the neutrophil-specific form of the low affinity F_c γreceptor III, CD16b, inhibited sGP binding specifically. Antibodies to CD62L, for example, did not inhibit sGP binding (Fig. 3). Binding to neutrophils correlated with their activation state and CD16b expression since no binding was observed in cells stimulated with phorbol 12-myristate 13-acetate (PMA) for 30 minutes, at which time CD16b expression was markedly decreased on these cells (Fig. 3, lower panel). Overexpression of this form

of CD16 on a heterologous cell type, 3T3 fibroblasts, did not confer sGP binding to these cells by FACS analysis, suggesting that CD16b is necessary but not sufficient for stable binding.

Binding of sGP did not inhibit neutrophil activation in response to potent pleiotropic activators (PMA, IL-8, or f-Met-Leu-Phe), as measured by down modulation of L-selectin expression using FACS analysis. In a defined serum-free medium, partial activation of neutrophils was observed, with a decrease in L-selectin expression at 4 hours (Fig. 4). Under these conditions, incubation of neutrophils with sGP supernatant prevented this decrease in L-selectin expression (Fig. 4). Because L-selectin was not required for sGP binding (Fig. 3), this effect was apparently indirect, through a mechanism not yet defined, possibly involving CD16b or carbohydrate interactions of the highly glycosylated sGP protein.

The expression of alternative Ebola virus glycoproteins in clinical infection has long been recognized, but their functional roles and cell specificity have not been defined. Early after infection, high levels of the secreted protein are found in the serum and precede fulminant replication and dissemination of virus systemically, at which time synthesis of transmembrane GP is markedly increased. Sanchez, A. et al., *PNAS (USA)* 93:3602 (1996). The inventors have now found that the binding specificities of these two molecules differ. It had been proposed that sGP may serve as a decoy to prevent recognition of GP, possibly to temporarily inhibit virus binding to target cells. The studies set forth herein suggest that this hypothesis is unlikely to be correct. The binding specificities of these proteins differ, and despite the fact that they are derived from the same viral gene, it has been surprisingly found that alternative forms of the glycoprotein have been selected for different functions.

Although these proteins share identical amino terminal sequences, their carboxyl terminal regions differ. Sanchez, A. et al., *Virus Res.* 29:215 (1993). These sequences are likely responsible for the differences in binding specificity, either through direct interactions in these domains or by their effect on multimerization. The secreted glycoprotein binds to neutrophils to prevent early events in activation, possibly serving to diminish any inflammatory responses which might provide innate immunity to the virus, facilitating productive viral replication. The subsequent increase in GP synthesis gives rise to virus which in turn could infect other cells. Filoviruses have been shown previously to infect and replicate in different cell types and appear to grow readily in endothelial cells *in vivo*. Peters, C.J. et al., *Fields Virology*, B.N. Fields, D.M. Knipe and P.M. Howley, Eds. (Lippincott-Raven, Philadelphia) (1996);

Schnittler, H.J. et al., *J. Clin. Invest.* 91:1301 (1993). The findings set forth herein suggest that its tropism for this cell type is probably determined by the specificity of GP. In Ebola infection, preferential binding and infection of microvascular endothelial cells may lead ultimately to a loss of capillary integrity that results in the severe hemorrhage observed in the terminal stages of this disease. The differential binding of these two gene products from the same viral structural gene generated by RNA editing suggests that they have evolved functionally to differentially affect immunity and infectivity. The ability to facilitate viral replication and target the virus to endothelial cells by alternative products of the same viral gene represents an efficient genetic mechanism which can account for different pathologic features of this disease. Inhibition of sGP binding to neutrophils and GP to endothelium is likely to ameliorate the effects of acute Ebola virus infection.

SPECIFIC EXAMPLE 2

I. Methods

Production of pseudotyped MuLV retroviruses expressing green fluorescent protein (GFP): 50% - 70% confluent 293 T cells in 60mm tissue culture dishes were transfected using the calcium phosphate method and the following plasmids: 0.3 μg 1012 GP(Z) (see Figure 8) or 1012 sGp-Gp(Z) (see Figure 9), 3 μg LZR-gfp, 2 μg pNGVL-gag-pol. After overnight transfection, fresh media was added to cells. Twenty hours later, the supernatants were harvested and filtered through a .45 μm filter.

Infection of HUVEC cells using the pseudotyped retroviruses: The day before infection, 30% - 50% confluent HUVEC cells were prepared in 6-well plates. 1 ml of pseudotyped retroviral supernatant was added to one well of the 6-well plates with 15 μg/ml of polybrene. Sixteen hours later, the viruses were removed and normal media was added. After 24 hours, the cells were lifted and GFP expression measured using FACS analysis.

Construction of 1012 sGP-GP(Z): 1012 sGP(Z) cells were digested with PstI and treated with Klenow, then digested with XbaI. 1012 GP(Z) cells were digested by EcoRI and treated with Klenow, then digested with KpnI. PstI/Klenow/XbaI treated sGP fragment and EcoRI/Klenow/KpnI treated GP fragment were then cloned into XbaI/KpnI treated pVR-1012 plasmid.

GP and sGP derivatives: The receptor recognition domain, mucin-like domain and/or TM domain of GP and sGP were mutated. The mutated GP and sGP was then tested for its ability to pseudotype and for cytotoxicity in producer cells.

II. Results

To determine the efficacy of targeting endothelium with the gene transfer vectors pseudotyped with GP of the present invention, HUVEC cells were infected with GP(Z) pseudotyped MuLV retrovirus (LZR-gfp) and sGP-GP(Z) pseudotyped MuLV retrovirus (LZR-gfp). Figures 5A-5C show the infection rate (GFP expression) measured using FACS analysis. As shown in Figure 5B, the GP(Z) pseudotyped MuLV retrovirus (LZR-gfp) was effective in targeting and expressing GFP in endothelium.

To determine whether mutating GP would effect its ability to pseudotype and/or decrease toxicity in producer cells, the receptor recognition domain, mucin-like domain and/or TM domain were mutated. Figure 10 shows the results. The optimal envelope is able to pseudotype but shows minimal toxicity.

The foregoing discussion discloses and describes merely exemplary embodiments of the present invention. One skilled in the art will readily recognize from such discussion, and from the accompanying drawings and claims, that various changes, modifications and variations can be made therein without departing from the spirit and scope of the invention as defined in the following claims.

All patents and other references cited herein are incorporated by reference as if fully set forth.

WE CLAIM:

1. A genetic construct comprising a gene operatively-linked to a carrier, wherein the carrier is associated with a transmembrane form of viral glycoprotein or derivative thereof.

2. The genetic construct of Claim 1, wherein the transmembrane form of viral glycoprotein or derivative thereof is expressed on the surface of the carrier.

3. The genetic construct of Claim 1, wherein the transmembrane form of viral glycoprotein or derivative thereof is from Ebola.

4. The genetic construct of Claim 1, wherein the carrier is a viral vector.

5. The genetic construct of Claim 1, wherein the carrier is a non-biologic gene targeting vehicle.

6. The genetic construct of Claim 4, wherein the viral vector is a retroviral vector.

7. The genetic construct of Claim 4, wherein the viral vector is a lentiviral vector.

8. The genetic construct of Claim 5, wherein the non-biologic gene targeting vehicle is a liposome.

9. The genetic construct of Claim 5, wherein the non-biologic gene targeting vehicle is a DNA-protein complex.

10. A method of targeting a gene to a cell comprising the step of administering to a cell population a genetic construct comprising the gene operatively-linked to a carrier, wherein the carrier is associated with a transmembrane form of viral glycoprotein or derivatives thereof.

11. The method of Claim 10, wherein the transmembrane form of viral glycoprotein or derivative thereof is expressed on the surface of the carrier.

12. The method of Claim 10, wherein the transmembrane form of viral glycoprotein or derivative thereof is from Ebola.

13. The method of Claim 10, wherein the carrier is a viral vector.

14. The method of Claim 10, wherein the step of administration is *ex vivo*.

15. The method of Claim 10, wherein the step of administration is *in vivo*.

16. The method of Claim 10, wherein the cell is an endothelial cell.

17. The method of Claim 10, wherein the cell is a hepatocyte.

18. The method of Claim 10, wherein the cell is a monocyte.

19. The method of Claim 10, wherein the cell is a dendritic cell.

20. The method of Claim 14, further comprising the step of introducing the cell population to a subject.

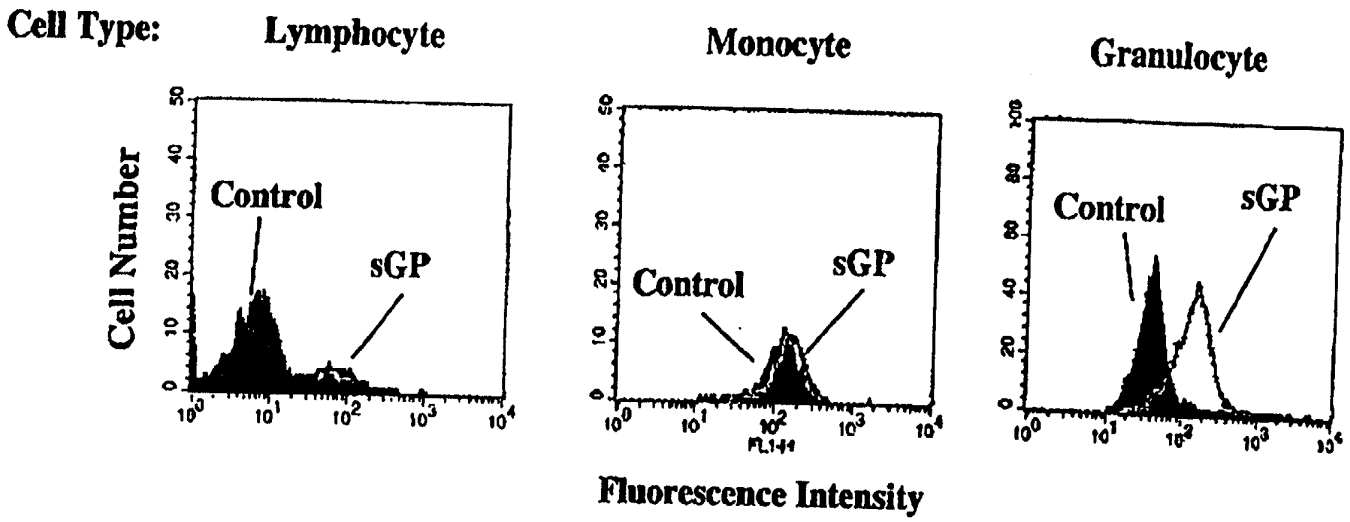


Figure 1A

Figure 1A1

Figure 1A2

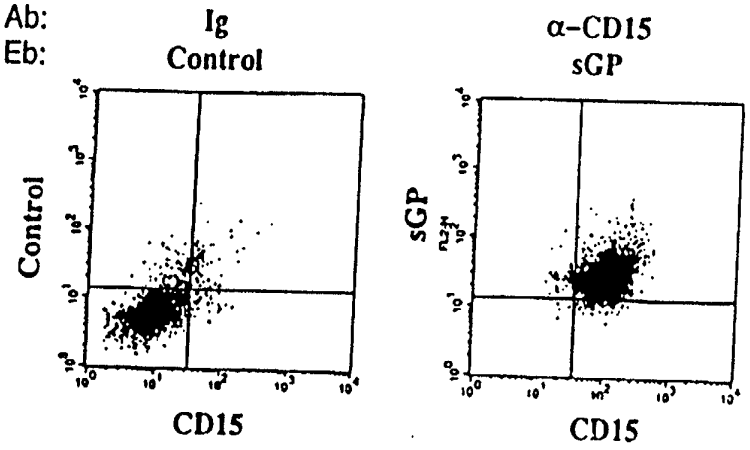


Figure 1B

Figure 1B1

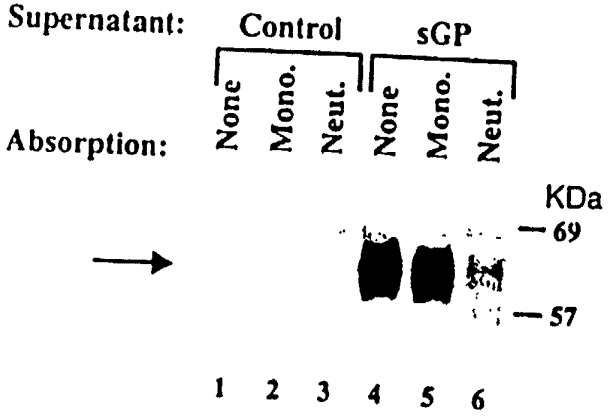


Figure 1C

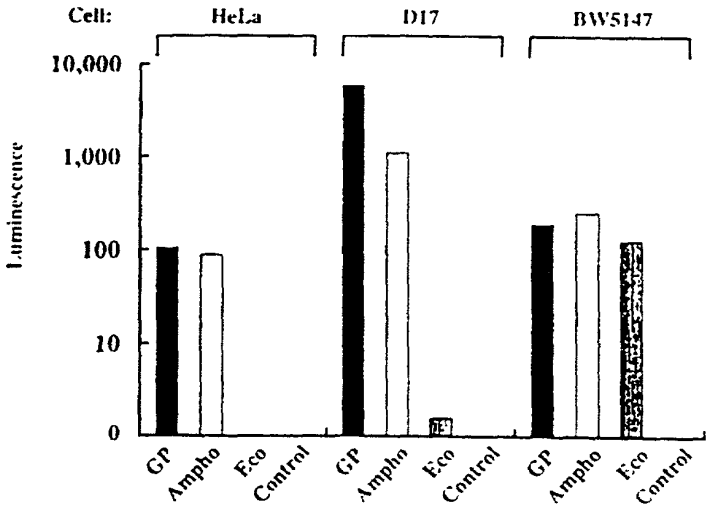


Figure 2A

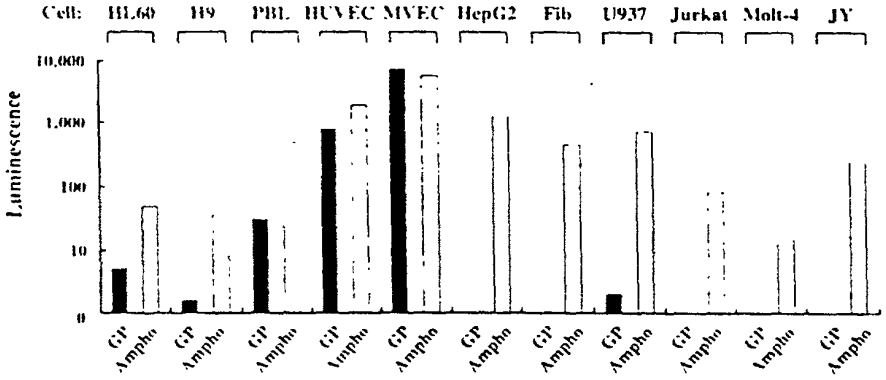


Figure 2B

Figure 2C

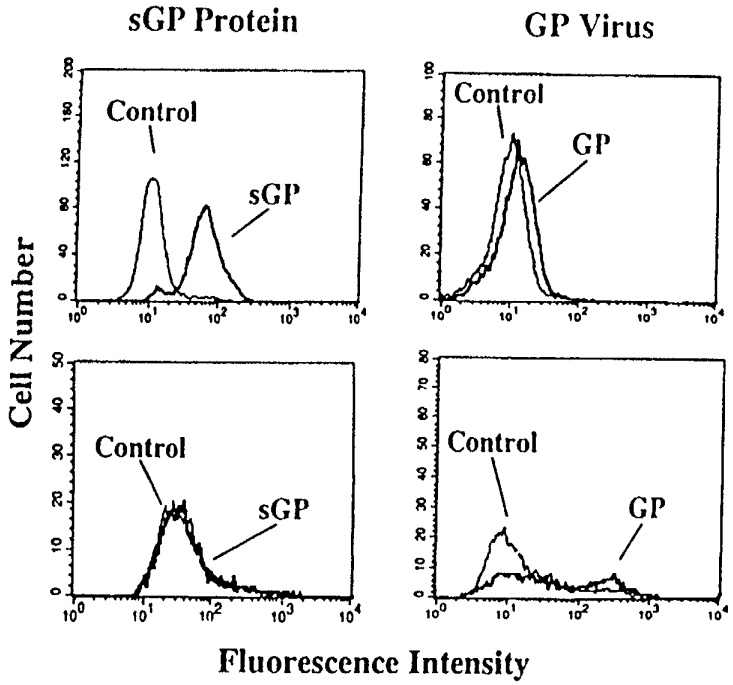


Figure 2C1

Figure 2C2

Figure 2C3

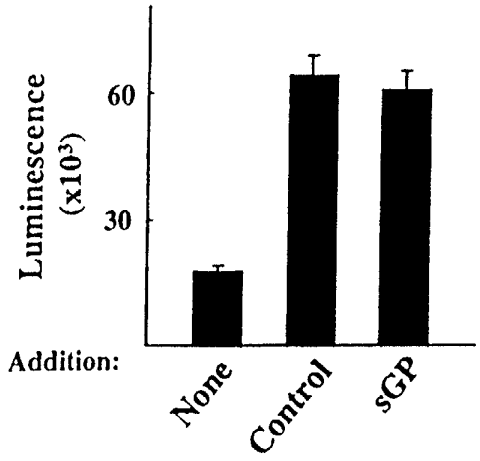


Figure 2D

Figure 3A

Ab: Control (Ig)

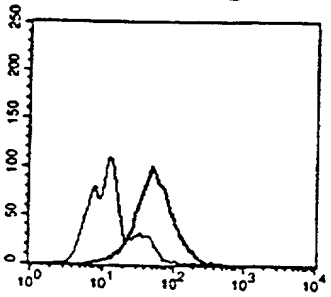


Figure 3B

α-CD16

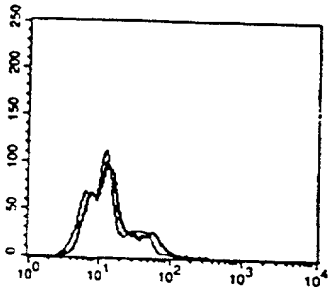


Figure 3C

Ab: Control (Ig)

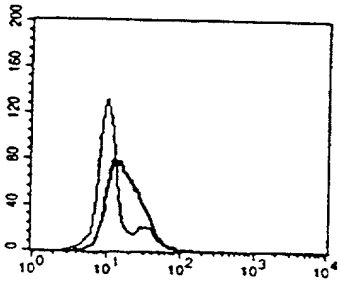
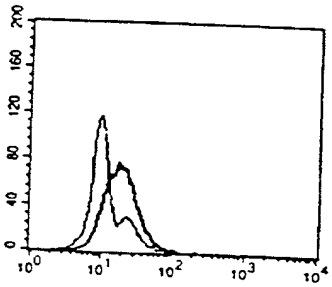


Figure 3D

α-CD62L



Cell Number

Figure 3E

Stim: None

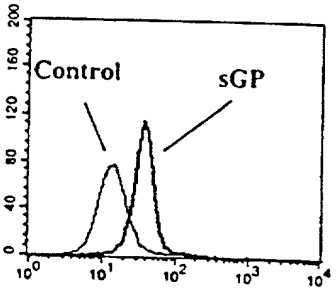
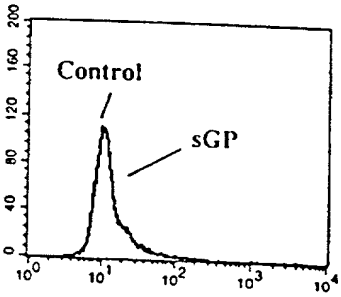


Figure 3F

PMA



Fluorescence Intensity

Figure 4A

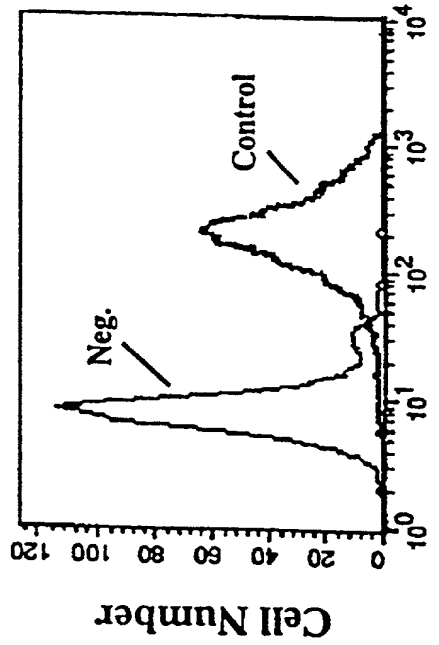
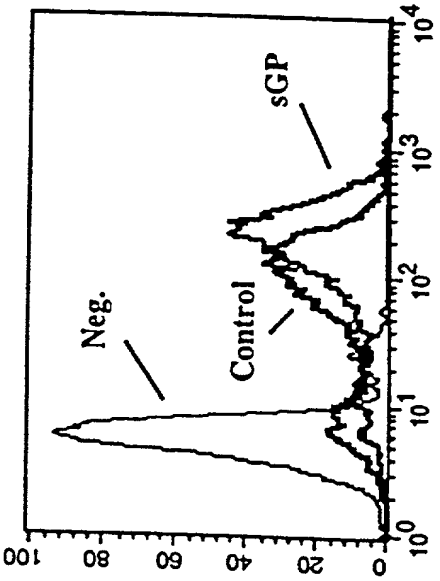


Figure 4B



Fluorescence Intensity

Figure 5A

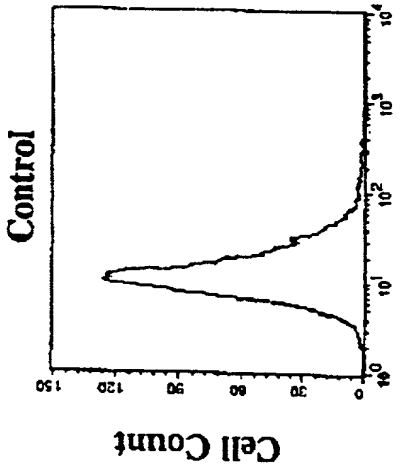


Figure 5B

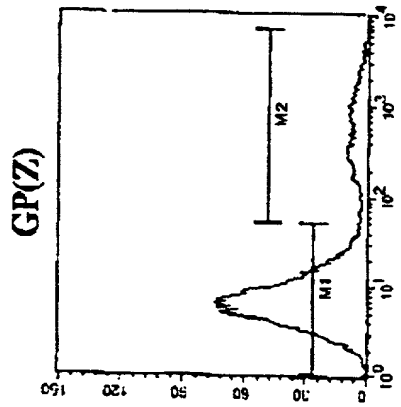
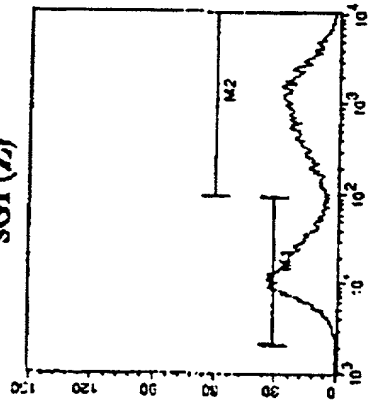


Figure 5C



Fluorescence Intensity

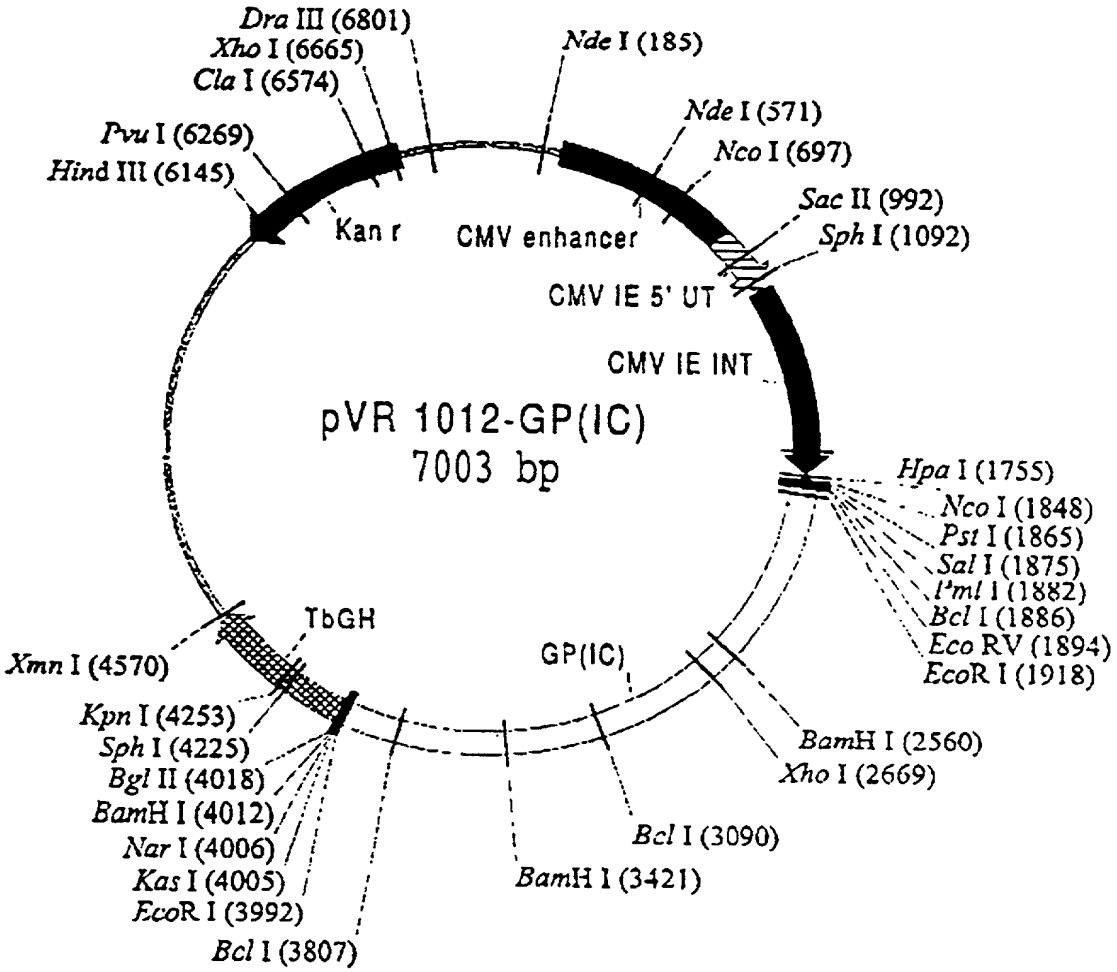


Figure 6

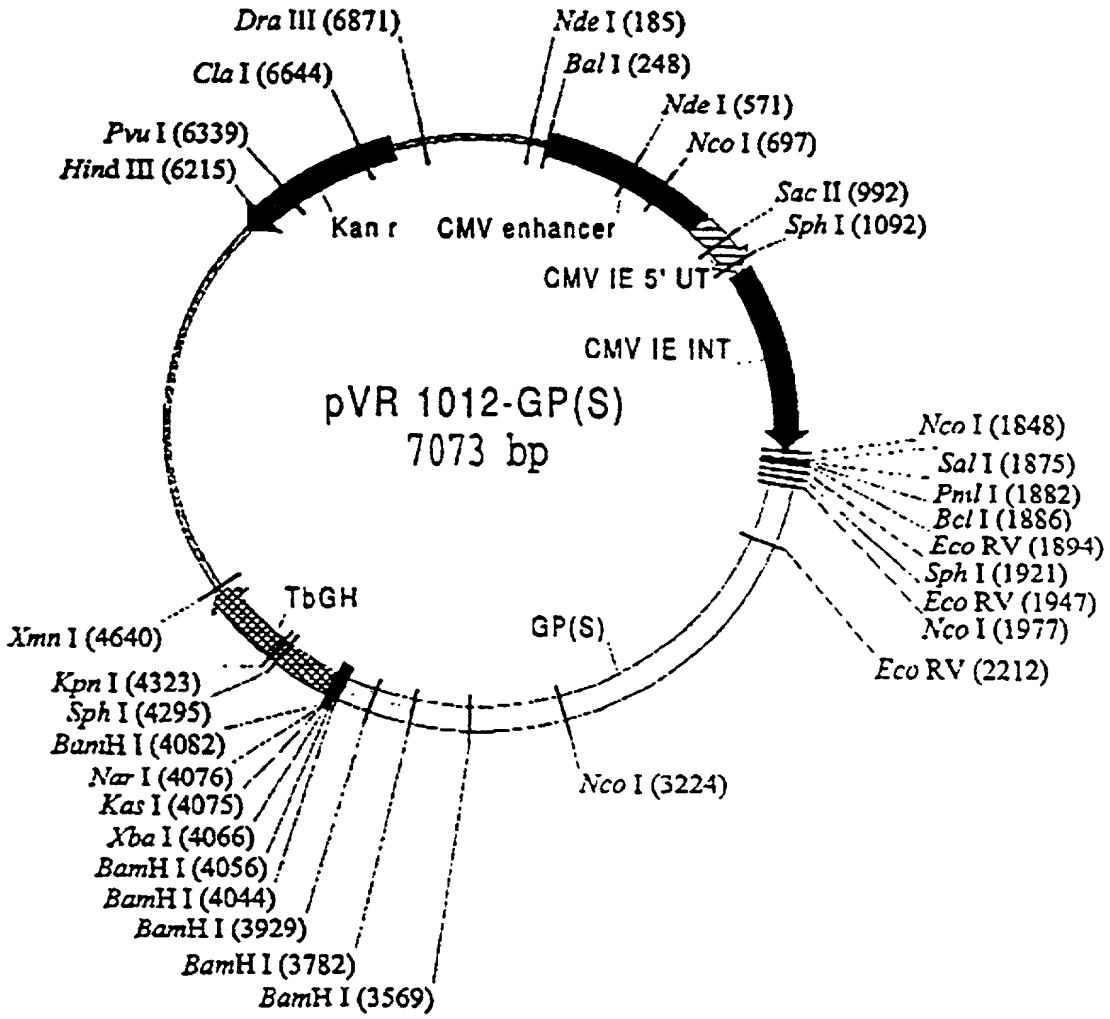

Dra III (6871)
Nde I (185)
Cla I (6644)
Bal I (248)
Nde I (571)
Pvu I (6339)
Nco I (697)
Hind III (6215)
Sac II (992)
Kan r
CMV enhancer
Sph I (1092)
CMV IE 5' UT
CMV IE INT
pVR 1012-GP(S)
7073 bp
Nco I (1848)
Sal I (1875)
Pml I (1882)
Bcl I (1886)
Eco RV (1894)
Sph I (1921)
Eco RV (1947)
Nco I (1977)
TbGH
Xmn I (4640)
GP(S)
Eco RV (2212)
Kpn I (4323)
Sph I (4295)
BamH I (4082)
Nar I (4076)
Kas I (4075)
Nco I (3224)
Xba I (4066)
BamH I (4056)
BamH I (4044)
BamH I (3929)
BamH I (3782)
BamH I (3569)

Figure 7

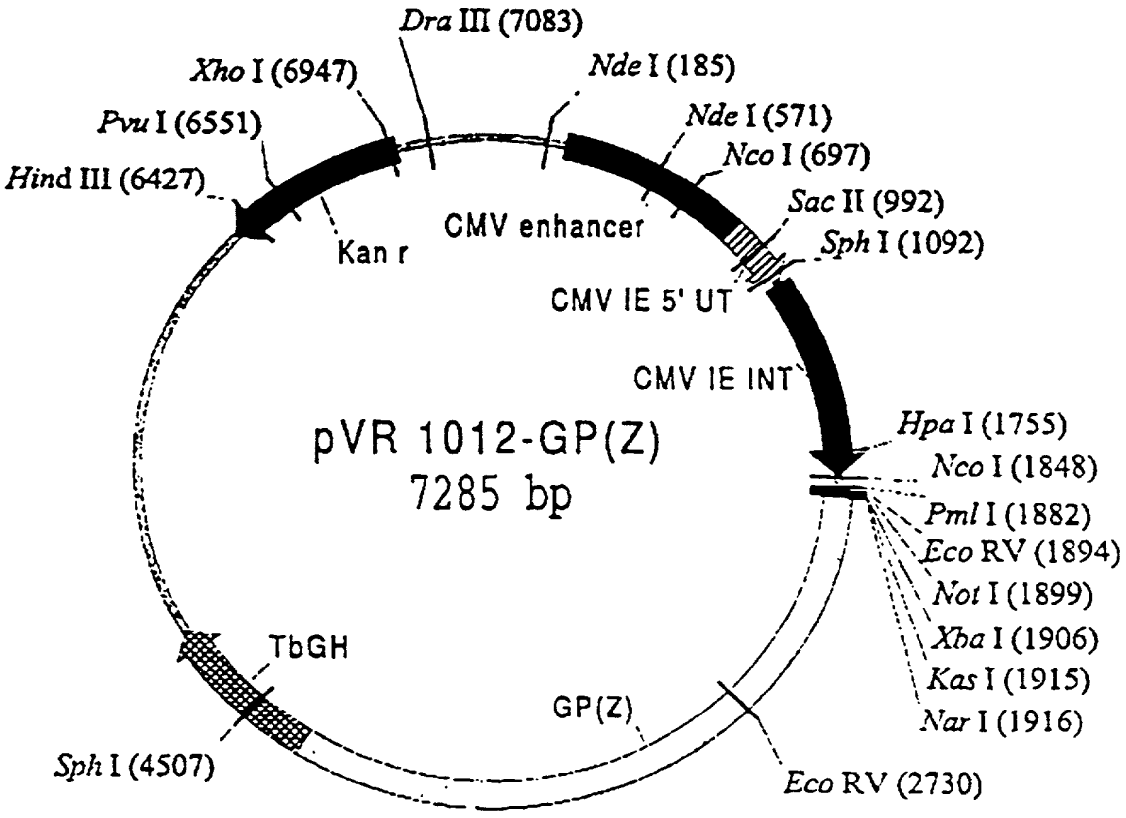


Figure 8

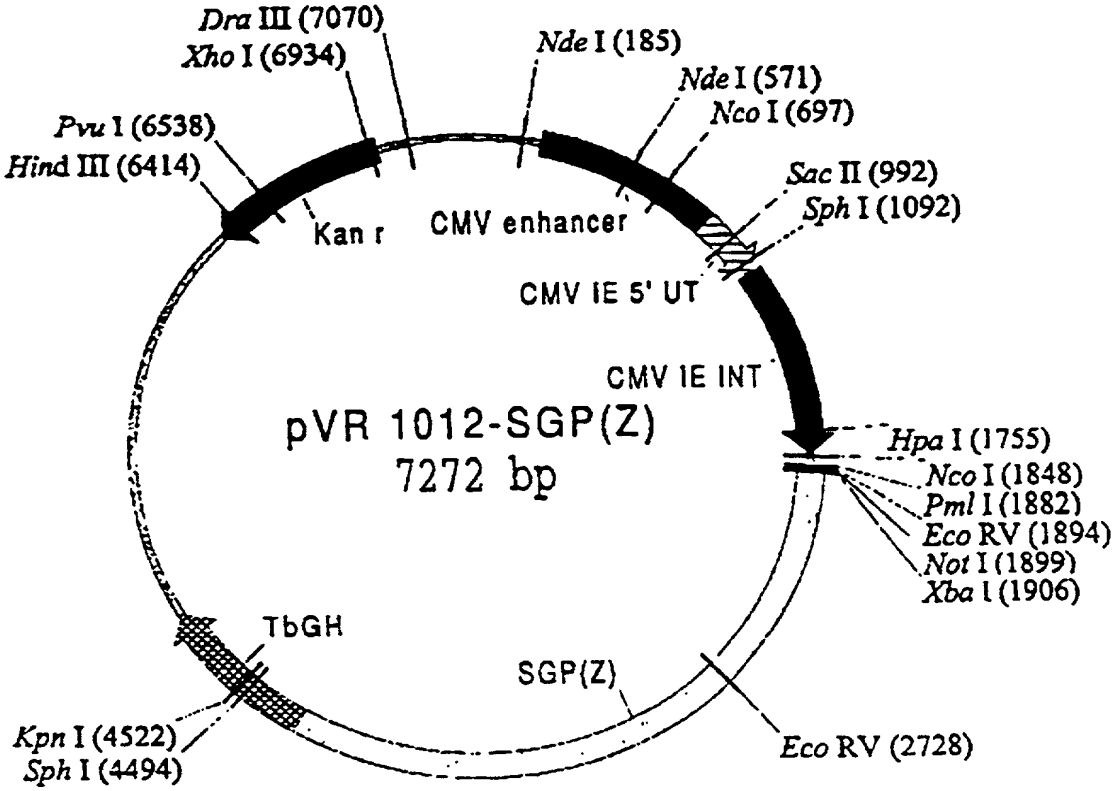


Figure 9

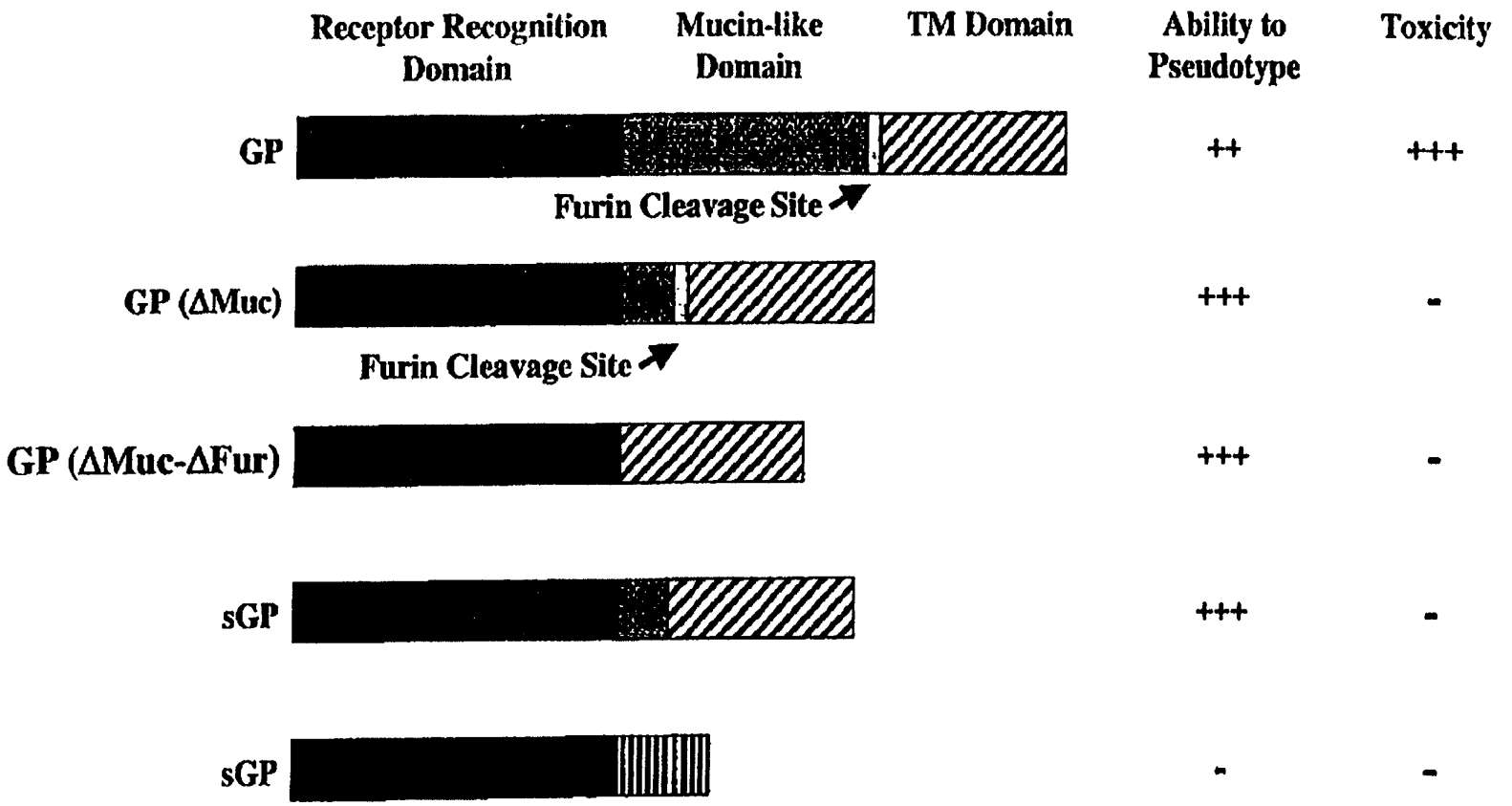


Figure 10

SEQUENCE LISTING ID NO: 1

pVR 1012-G2(IC)

General Description

DNA pVR 1012-G2(IC)
Local object
Created: 09/14/98 04:17PM
Last Modification Date: ? (no data)
length: 7003 bp
storage type: Basic
form: Circular

Comments

Restriction Map

BglII: 1 site AGATCT / TCTAGA

ClaI: 1 site ATCGAT / TAGCTA

DraIII: 1 site CACNNNGTC / GTGNNNCAC

EcoRV: 1 site GATATC / CTATAG

HindIII: 1 site AAGCTT / TTCGAA

HpaI: 1 site GTTAAC / CAATTG

KasI: 1 site GGCGCC / CCGCGG

KpnI: 1 site GGTACC / CCATGG

NarI: 1 site GGCGCC / CCGCGG

PmlI: 1 site CACGTG / GTGCAC

PstI: 1 site CTGCAG / GACGTC

PvuI: 1 site CGATCG / GCTAGC

SacII: 1 site CCGCGG / GGCGCC

SalI: 1 site GTCGAC / CAGCTG

XmnI: 1 site GAANNNNTTC / CTTNNNNAAG

EcoRI: 2 sites GAATTC / CTTAAG

NcoI: 2 sites CCATGG / GGTACC

NdeI: 2 sites CATATG / GTATAC

SphI: 2 sites GCATGC / CGTACG

XhoI: 2 sites CTCGAG / GAGCTC

BamHI: 3 sites GGATCC / CCTAGG

BclI: 3 sites TGATCA ACTAGT

Functional Map

CDS (4 signals)

CMV IE 5' UT
Start: 886 End: 1129

CMV IE INT
Start: 1130 End: 1840

TbGH
Start: 4020 End: 4572

Kan r
Start: 6068 End: 6690 (Complementary)

Misc_feature (2 signals)

CMV enhancer
Start: 248 End: 885

GP(IC)
Start: 1870 End: 4019

Annotations

```
   1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

  51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
      CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

 101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
      AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                          NdeI
 151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
      GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACGC CACACTTTAT

 201  CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
      GGCGTGTCTA CGCATTCCTC TTTTATGGCG TAGTCTAACC GATAACCGGT

 251  TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
      AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

 301  TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
      AGGTTGTAAT GGCGGTACAA CTGTAACTAA TAACTGATCA ATAATTATCA

 351  AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
      TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA

 401  ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG
      TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC

 451  CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
      GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTGC GGTTATCCCT

 501  CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
      GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTG ACGGGTGAAC

                      NdeI
 551  GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
      CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT

 601  TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
      ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGGAATACCC

                                                      NcoI
 651  ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
      TGAAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

      NcoI
 701  GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
      CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG

 751  ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
      TGCCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAAA

 801  GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
      CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT
```

```
 851  TTGACGCAAA TGGGCGGTAG GCGTGTACGG TGGGAGGTCT ATATAAGCAG
      AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC
 ..........................................................................

 901  AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
      TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA
 ..........................................................................

                                                  SacII
                                                  ------

 951  TTCACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
      AAGTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT
 ..........................................................................

1001  CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
      GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG
 ..........................................................................

                                                  SphI
                                                  ----

1051  CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
      GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA
 ..........................................................................

1101  TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
      AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC
 ..........................................................................

1151  TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
      ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG
 ..........................................................................

1201  TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
      ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT
 ..........................................................................

1251  CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
      GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG
 ..........................................................................

1301  ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
      TGCCTGAGAC ATAAAAATGT CCTACCCCAG GGTAAATAAT AAATGTTTAA
 ..........................................................................

1351  CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
      GTGTATATGT TGTTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT
 ..........................................................................

1401  GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
      CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA
 ..........................................................................

1451  TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
      AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG
 ..........................................................................

1501  AGCGGCTCAT GGTCGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
      TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC
 ..........................................................................

1551  ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
      TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC
 ..........................................................................

1601  TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
      ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGAGCGTGC
 ..........................................................................

1651  CCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
      GGACTGCGTC TACCTTCTGA ATTCCGTCGC CGTCTTCTTC TACGTCCGTC
 ..........................................................................

1701  CCGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
      GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG
 ..........................................................................
```

```
        HpaI
1751  TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
      ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC
..........................................................................
                                                    NcoI
1801  CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
      GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA
..........................................................................
                            SalI
      NcoI      PstI                  PmlI   BclI   EcoRV
1851  GGGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
      CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC
..........................................................................
                     EcoRI
1901  CCGCGCGGCC GCTCTAGAAT TCTCTAATCA CAGTCATCAT GGGAGCGTCA
      GGCGCGCCGG CGAGATCTTA AGAGATTAGT GTCAGTAGTA CCCTCGCAGT
..........................................................................
1951  GGGATTCTGC AATTGCCCCG TGAGCGCTTC AGGAAAACAT CTTTCTTTGT
      CCCTAAGACG TTAACGGGGC ACTCGCGAAG TCCTTTTGTA GAAAGAAACA
..........................................................................
2001  TTGGGTAATA ATCCTATTCC ATAAAGTCTT TTCAATCCCG TTGGGGGTTG
      AACCCATTAT TAGGATAAGG TATTTCAGAA AAGTTAGGGC AACCCCCAAC
..........................................................................
2051  TACACAACAA TACCCTACAA GTGAGTGATA TTGACAAGTT TGTGTGCCGA
      ATGTGTTGTT ATGGGATGTT CACTCACTAT AACTGTTCAA ACACACGGCT
..........................................................................
2101  GACAAACTCT CTTCAACTAG CCAATTGAAG TCAGTCGGGT TGAACTTGGA
      CTGTTTGAGA GAAGTTGATC GGTTAACTTC AGTCAGCCCA ACTTGAACCT
..........................................................................
2151  GGGCAATGGA GTAGCAACTG ATGTACCAAC GGCAACCAAA AGATGGGGTT
      CCCGTTACCT CATCGTTGAC TACATGGTTG CCGTTGGTTT TCTACCCCAA
..........................................................................
2201  TTCGAGCTGG TGTTCCACCA AAGGTGGTAA ATTACGAAGC TGGAGAATGG
      AAGCTCGACC ACAAGGTGGT TTCCACCATT TAATGCTTCG ACCTCTTACC
..........................................................................
2251  GCTGAGAACT GTTATAACCT GGCTATAAAG AAAGTTGATG GTAGTGAGTG
      CGACTCTTGA CAATATTGGA CCGATATTTC TTTCAACTAC CATCACTCAC
..........................................................................
2301  CCTACCAGAA GCCCCTGAGG GAGTGAGGGA TTTTCCCCGT TGCCGCTATG
      GGATGGTCTT CGGGGACTCC CTCACTCCCT AAAAGGGGCA ACGGCGATAC
..........................................................................
2351  TACACAAAGT CTCAGGAACT GGACCATGCC CAGGAGGACT CGCCTTTCAC
      ATGTGTTTCA GAGTCCTTGA CCTGGTACGG GTCCTCCTGA GCGGAAAGTG
..........................................................................
2401  AAAGAAGGAG CCTTCTTCCT GTATGACCGA CTCGCATCAA CAATCATTTA
      TTTCTTCCTC GGAAGAAGGA CATACTGGCT GAGCGTAGTT GTTAGTAAAT
..........................................................................
2451  TCGGGGTACA ACCTTTGCCG AAGGAGTTAT TGCATTTCTG ATCTTGCCTA
      AGCCCCATGT TGGAAACGGC TTCCTCAATA ACGTAAAGAC TAGAACGGAT
..........................................................................
2501  AGGCGCGAAA GGATTTTTTC CAGTCTCCTC CATTGCATGA GCCTGCCAAC
      TCCGCGCTTT CCTAAAAAAG GTCAGAGGAG GTAACGTACT CGGACGGTTG
..........................................................................
```

BamHI

```
2551  ATGACCACGG ATCCCTCCAG TTACTATCAC ACGACAACAA TAAACTACGT
      TACTGGTGCC TAGGGAGGTC AATGATAGTG TGCTGTTGTT ATTTGATGCA
```

```
2601  GGTTGATAAT TTTGGAACCA ACACCACAGA GTTTCTGTTC CAAGTCGATC
      CCAACTATTA AAACCTTGGT TGTGGTGTCT CAAAGACAAG GTTCAGCTAG
```

XhoI

```
2651  ATTTGACGTA TGTGCAGCTC GAGGCAAGAT TCACACCACA ATTCCTTGTC
      TAAACTGCAT ACACGTCGAG CTCCGTTCTA AGTGTGGTGT TAAGGAACAG
```

```
2701  CTCCTAAATG AAACCATCTA CTCTGATAAC CGCAGAAGTA ACACAACAGG
      GAGGATTTAC TTTGGTAGAT GAGACTATTG GCGTCTTCAT TGTGTTGTCC
```

```
2751  AAAACTAATC TGGAAAATAA ATCCCACTGT TGATACCAGC ATGGGTGAGT
      TTTTGATTAG ACCTTTTATT TAGGGTGACA ACTATGGTCG TACCCACTCA
```

```
2801  GGGCTTTCTG GGAAAATAAA AAAACTCCAC AAAAACCCTT TCAAGTGAAG
      CCCGAAAGAC CCTTTTATTT TTTTGAAGTG TTTTTGGGAA AGTTCACTTC
```

```
2851  AGTTGTCTTT CGTACCTGTA CCAGAAACCC AGAACCAGGT CCTTGACACG
      TCAACAGAAA GCATGGACAT GGTCTTTGGG TCTTGGTCCA GGAACTGTGC
```

```
2901  ACAGCGACGG TCTCTCCTCC CATCTCCGCC CACAACCACG CAGGCGAAGA
      TGTCGCTGCC AGAGAGGAGG GTAGAGGCGG GTGTTGGTGC GTCCGCTTCT
```

```
2951  CCACAAAGAA TTGGTTTCAG AGGATTCCAC TCCAGTGGTT CAGATGCAAA
      GGTGTTTCTT AACCAAAGTC TCCTAAGGTG AGGTCACCAA GTCTACGTTT
```

```
3001  ACATCAACGG AAAGGACACA ATGCCAACCA CAGTGACGGG TGTACCAACA
      TGTAGTTCCC TTTCCTGTGT TACGGTTGGT GTCACTGCCC ACATGGTTGT
```

BclI

```
3051  ACCACACCCT CTCCATTTCC AATCAATGCT CGCAACACTG ATCATACCAA
      TGGTGTGGGA GAGGTAAAGG TTAGTTACGA GCGTTGTGAC TAGTATGGTT
```

```
3101  ATCATTTATC GGCCTGGAGG GGCCCCAAGA AGACCACAGC ACCACACAGC
      TAGTAAATAG CCGGACCTCC CCGGGGTTCT TCTGGTGTCG TGGTGTGTCG
```

```
3151  CTGCCAAGAC CACCAGCCAA CCAACCAACA GCACAGAATC GACGACACTA
      GACGGTTCTG GTGGTCGGTT GGTTGGTTGT CGTGTCTTAG CTGCTGTGAT
```

```
3201  AACCCAACAT CAGAGCCCTC CAGTAGAGGC ACGGGACCAT CCAGCCCCAC
      TTGGGTTGTA GTCTCGGGAG GTCATCTCCG TGCCCTGGTA GGTCGGGGTG
```

```
3251  GGTCCCCAAC ACCACAGAAA GCCACGCCGA ACTTGGCAAG ACAACCCCAA
      CCAGGGGTTG TGGTGTCTTT CGGTGCGGCT TGAACCGTTC TGTTGGGGTT
```

```
3301  CCACACTCCC AGAACAGCAC ACTGCCGCCA GTGCCATTCC AAGAGCCGTG
      GGTGTGAGGG TCTTGTCGTG TGACGGCGGT CACGGTAAGG TTCTCGGCAC
```

```
3351  CACCCCGACG AACTCAGTGG ACCTGGCTTC CTGACGAACA CAATACGGGG
      GTGGGGCTGC TTGAGTCACC TGGACCGAAG GACTGCTTGT GTTATGCCCC
```

```
                          BamHI
                          ------
3401  GGTGACAAAT CTCCTGACAG GATCCAGAAG AAAGCGAAGG GATGTCACTC
      CCACTGTTTA GAGGACTGTC CTAGGTCTTC TTTCGCTTCC CTACAGTGAG
 ----------------------------------------------------------------------

3451  CCAATACACA ACCCAAATGC AACCCAAACC TGCACTATTG GACAGCCTTG
      GGTTATGTGT TGGGTTTACG TTGGGTTTGG ACGTGATAAC CTGTCGGAAC
 ----------------------------------------------------------------------

3501  GATGAGGGTG CTGCCATAGG TTTAGCCTGG ATACCATACT TCGGGCCAGC
      CTACTCCCAC GACGGTATCC AAATCGGACC TATGGTATGA AGCCCGGTCG
 ----------------------------------------------------------------------

3551  AGCTGAGGGA ATTTACACTG AAGGCATAAT CGAGAATCAA AATGGATTGA
      TCGACTCCCT TAAATGTGAC TTCCGTATTA GCTCTTAGTT TTACCTAACT
 ----------------------------------------------------------------------

3601  TCTGTGGATT GAGGCAGCTG GCCAACGAAA CGACACAAGC TCTTCAATTG
      AGACACCTAA CTCCGTCGAC CGGTTGCTTT GCTGTGTTCG AGAAGTTAAC
 ----------------------------------------------------------------------

3651  TTCTTAAGGG CAACTACTGA GTTGCGTACA TTCTCTATAC TAAATCGGAA
      AAGAATTCCC GTTGATGACT CAACGCATGT AAGAGATATG ATTTAGCCTT
 ----------------------------------------------------------------------

3701  AGCAATAGAC TTCTTGCTCC AAAGATGGGG AGGAACATGT CACATTCTAG
      TCGTTATCTG AAGAACGAGG TTTCTACCCC TCCTTGTACA GTGTAAGATC
 ----------------------------------------------------------------------

3751  GGCCTGATTG TTGCATTGAA CCCCAAGATT GGACCAAAAA TATCACTGAT
      CCGGACTAAC AACGTAACTT GGGGTTCTAA CCTGGTTTTT ATAGTGACTA
 ----------------------------------------------------------------------

          BclI
          ----
3801  AAAATTGATC AAATAATCCA TGACTTTGTC GATAATAATC TTCCAAATCA
      TTTTAACTAG TTTATTAGGT ACTGAAACAG CTATTATTAG AAGGTTTAGT
 ----------------------------------------------------------------------

3851  GAATGATGGC AGCAACTGGT GGACTGGATG GAAACAATGG GTTCCTGCTG
      CTTACTACCG TCGTTGACCA CCTGACCTAC CTTTGTTACC CAAGGACGAC
 ----------------------------------------------------------------------

3901  GAATAGGAAT CACAGGAGTA ATCATTGCTA TTATTGCTTT GCTGTGCATT
      CTTATCCTTA GTGTCCTCAT TAGTAACGAT AATAACGAAA CGACACGTAA
 ----------------------------------------------------------------------

                                                    EcoRI
                                                    -----
3951  TGCAAATTCA TGCTTTGAAC TAATATAGCA TCATACTTTA GAATTCTAGA
      ACGTTTAAGT ACGAAACTTG ATTATATCGT AGTATGAAAT CTTAAGATCT
 ----------------------------------------------------------------------

         NarI
         ----
         XasI    BamHI BglII
         ----    -----------
4001  CCAGGCGCCT GGATCCAGAT CTGCTGTGCC TTCTAGTTGC CAGCCATCTG
      GGTCCGCGGA CCTAGGTCTA GACGACACGG AAGATCAACG GTCGGTAGAC
 ----------------------------------------------------------------------

4051  TTGTTTGCCC CTCCCCCGTG CCTTCCTTGA CCCTGGAAGG TGCCACTCCC
      AACAAACGGG GAGGGGGCAC GGAAGGAACT GGGACCTTCC ACGGTGAGGG
 ----------------------------------------------------------------------

4101  ACTGTCCTTT CCTAATAAAA TGAGGAAATT GCATCGCATT GTCTGAGTAG
      TGACAGGAAA GGATTATTTT ACTCCTTTAA CGTAGCGTAA CAGACTCATC
 ----------------------------------------------------------------------

4151  GTGTCATTCT ATTCTGGGGG GTGGGGTGGG GCAGCACAGC AAGGGGGAGG
      CACAGTAAGA TAAGACCCCC CACCCCACCC CGTCGTGTCG TTCCCCCTCC
 ----------------------------------------------------------------------
```

```
                               SphI                          KpnI
                            ---------                         ---
4201  ATTGGGAAGA CAATAGCAGG CATGCTGGGG ATGCGGTGGG CTCTATGGGT
      TAACCCTTCT GTTATCGTCC GTACGACCCC TACGCCACCC GAGATACCCA
...........................................................................
      KpnI
      ---
4251  ACCCAGGTGC TGAAGAATTG ACCCGGTTCC TCCTGGGCCA GAAAGAAGCA
      TGGGTCCACG ACTTCTTAAC TGGGCCAAGG AGGACCCGGT CTTTCTTCGT
...........................................................................
4301  GGCACATCCC CTTCTCTGTG ACACACCCTG TCCACGCCCC TGGTTCTTAG
      CCGTGTAGGG GAAGAGACAC TGTGTGGGAC AGGTGCGGGG ACCAAGAATC
...........................................................................
4351  TTCCAGCCCC ACTCATAGGA CACTCATAGC TCAGGAGGGC TCCGCCTTCA
      AAGGTCGGGG TGAGTATCCT GTGAGTATCG AGTCCTCCCG AGGCGGAAGT
...........................................................................
4401  ATCCCACCCG CTAAAGTACT TGGAGCGGTC TCTCCCTCCC TCATCAGCCC
      TAGGGTGGGC GATTTCATGA ACCTCGCCAG AGAGGGAGGG AGTAGTCGGG
...........................................................................
4451  ACCAAACCAA ACCTAGCCTC CAAGAGTGGG AAGAAATTAA AGCAAGATAG
      TGGTTTGGTT TGGATCGGAG GTTCTCACCC TTCTTTAATT TCGTTCTATC
...........................................................................
4501  GCTATTAAGT GCAGAGGGAG AGAAAATGCC TCCAACATGT GAGGAAGTAA
      CGATAATTCA CGTCTCCCTC TCTTTTACGG AGGTTGTACA CTCCTTCATT
...........................................................................
                          XmnI
                      ------------
4551  TGAGAGAAAT CATAGAATTT CTTCCGCTTC CTCGCTCACT GACTCGCTGC
      ACTCTCTTTA GTATCTTAAA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG
...........................................................................
4601  GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA
      CGAGCCAGCA AGCCGACGCC GCTCGCCATA GTCGAGTGAG TTTCCGCCAT
...........................................................................
4651  ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC
      TATGCCAATA GGTGTCTTAG TCCCCTATTG CGTCCTTTCT TGTACACTCG
...........................................................................
4701  AAAAGGCCAG CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT
      TTTTCCGGTC GTTTTCCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA
...........................................................................
4751  TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA
      AAAAGGTATC CGAGGCGGGG GGACTGCTCG TAGTGTTTTT AGCTGCGAGT
...........................................................................
4801  AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC
      TCAGTCTCCA CCGCTTTGGG CTGTCCTGAT ATTTCTATGG TCCGCAAAGG
...........................................................................
4851  CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG
      GGGACCTTCG AGGGAGCACG CGAGAGGACA AGGCTGGGAC GGCGAATGGC
...........................................................................
4901  GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC
      CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG
...........................................................................
4951  TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG
      AGTGCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC
...........................................................................
5001  CTGTGTGCAC GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA
      GACACACGTG CTTGGGGGGC AAGTCGGGCT GGCGACGCGG AATAGGCCAT
...........................................................................
```

```
5051  ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA
      TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CGGTGACCGT
 ..........................................................................

5101  GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC
      CGTCGGTGAC CATTGTCCTA ATCGTCTCGC TCCATACATC CGCCACGATG
 ..........................................................................

5151  AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA AGGACAGTAT
      TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT TCCTGTCATA
 ..........................................................................

5201  TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT
      AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA
 ..........................................................................

5251  AGCTCTTGAT CCGGCAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT
      TCGAGAACTA GGCCGTTTGT TTGGTGGCGA CCATCGCCAC CAAAAAACA
 ..........................................................................

5301  TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
      AACGTTCGTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTTCTAGGAA
 ..........................................................................

5351  TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA
      ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT
 ..........................................................................

5401  GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT
      CCCTAAAACC AGTACTCTAA TAGTTTTTCC TAGAAGTGGA TCTAGGAAAA
 ..........................................................................

5451  AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT
      TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA
 ..........................................................................

5501  GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC
      CCAGACTGTC AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG
 ..........................................................................

5551  TGTCTATTTC GTTCATCCAT AGTTGCCTGA CTCCGGGGGG GGGGGGCGCT
      ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGCCCCCC CCCCCCGCGA
 ..........................................................................

5601  GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG ACTCATACCA GGCCTGAATC
      CTCCAGACGG AGCACTTCTT CCACAACGAC TGAGTATGGT CCGGACTTAG
 ..........................................................................

5651  GCCCCATCAT CCAGCCAGAA AGTGAGGGAG CCACGGTTGA TGAGAGCTTT
      CGGGGTAGTA GGTCGGTCTT TCACTCCCTC GGTGCCAACT ACTCTCGAAA
 ..........................................................................

5701  GTTGTAGGTG GACCAGTTGG TGATTTTGAA CTTTTGCTTT GCCACGGAAC
      CAACATCCAC CTGGTCAACC ACTAAAACTT GAAAACGAAA CGGTGCCTTG
 ..........................................................................

5751  GGTCTGCGTT GTCGGGAAGA TGCGTGATCT GATCCTTCAA CTCAGCAAAA
      CCAGACGCAA CAGCCCTTCT ACGCACTAGA CTAGGAAGTT GAGTCGTTTT
 ..........................................................................

5801  GTTCGATTTA TTCAACAAAG CCGCCGTCCC GTCAAGTCAG CGTAATGCTC
      CAACCTAAAT AAGTTGTTTC GGCGGCAGGG CAGTTCAGTC GCATTACGAG
 ..........................................................................

5851  TGCCAGTGTT ACAACCAATT AACCAATTCT GATTAGAAAA ACTCATCGAG
      ACGGTCACAA TGTTGGTTAA TTGGTTAAGA CTAATCTTTT TGAGTAGCTC
 ..........................................................................

5901  CATCAAATGA AACTGCAATT TATTCATATC AGGATTATCA ATACCATATT
      GTAGTTTACT TTGACGTTAA ATAAGTATAG TCCTAATAGT TATGGTATAA
 ..........................................................................

5951  TTTGAAAAAG CCGTTTCTGT AATGAAGGAG AAAACTCACC GAGGCAGTTC
      AAACTTTTTC GGCAAAGACA TTACTTCCTC TTTTGAGTGG CTCCGTCAAG
 ..........................................................................
```

```
6001  CATAGGATGG CAAGATCCTG GTATCGGTCT GCGATTCCGA CTCGTCCAAC
      GTATCCTACC GTTCTAGGAC CATAGCCAGA CGCTAAGGCT GAGCAGGTTG
 ..........................................................................

6051  ATCAATACAA CCTATTAATT TCCCCTCGTC AAAAATAAGG TTATCAAGTG
      TAGTTATGTT GGATAATTAA AGGGGAGCAG TTTTTATTCC AATAGTTCAC
 ..........................................................................

                                                      HindIII
                                                      ------
6101  AGAAATCACC ATGAGTGACG ACTGAATCCG GTGAGAATGG CAAAAGCTTA
      TCTTTAGTGG TACTCACTGC TGACTTAGGC CACTCTTACC GTTTTCGAAT
 ..........................................................................

6151  TGCATTTCTT TCCAGACTTG TTCAACAGGC CAGCCATTAC GCTCGTCATC
      ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG GTCGGTAATG CGAGCAGTAG
 ..........................................................................

6201  AAAATCACTC GCATCAACCA AACCGTTATT CATTCGTGAT TGCGCCTGAG
      TTTTAGTGAG CGTAGTTGGT TTGGCAATAA GTAAGCACTA ACGCGGACTC
 ..........................................................................

                      PvuI
                      ----
6251  CGAGACGAAA TACGCGATCG CTGTTAAAAG GACAATTACA AACAGGAATC
      GCTCTGCTTT ATGCGCTAGC GACAATTTTC CTGTTAATGT TTGTCCTTAG
 ..........................................................................

6301  GAATGCAACC GGCGCAGGAA CACTGCCAGC GCATCAACAA TATTTTCACC
      CTTACGTTGG CCGCGTCCTT GTGACGGTCG CGTAGTTGTT ATAAAAGTGG
 ..........................................................................

6351  TGAATCAGGA TATTCTTCTA TACCTGGAA TGCTGTTTTC CCGGGGATCG
      ACTTAGTCCT ATAAGAAGAT TATGGACCTT ACGACAAAAG GGCCCCTAGC
 ..........................................................................

6401  CAGTCGTGAG TAACCATGCA TCATCAGGAG TACGGATAAA ATGCTTGATG
      GTCAGCACTC ATTGGTACGT AGTAGTCCTC ATGCCTATTT TACGAACTAC
 ..........................................................................

6451  GTCGGAAGAG GCATAAATTC CGTCAGCCAG TTTAGTCTGA CCATCTCATC
      CAGCCTTCTC CGTATTTAAG GCAGTCGGTC AAATCAGACT GGTAGAGTAG
 ..........................................................................

6501  TGTAACATCA TTGGCAACGC TACCTTTGCC ATGTTTCAGA AACAACTCTG
      ACATTGTAGT AACCGTTGCG ATGGAAACGG TACAAAGTCT TTGTTGAGAC
 ..........................................................................

                               ClaI
                               ----
6551  GCGCATCGGG CTTCCCATAC AATCGATAGA TTGTCGCACC TGATTGCCCG
      CCCGTAGCCC GAAGGGTATG TTAGCTATCT AACAGCGTGG ACTAACGGGC
 ..........................................................................

6601  ACATTATCGC GAGCCCATTT ATACCCATAT AAATCAGCAT CCATGTTGGA
      TGTAATAGCG CTCGGGTAAA TATGGGTATA TTTAGTCGTA GGTACAACCT
 ..........................................................................

                      XhoI
                      ----
6651  ATTTAATCGC GGCCTCGAGC AAGACGTTTC CCGTTGAATA TGGCTCATAA
      TAAATTAGCG CCGGAGCTCG TTCTGCAAAG GGCAACTTAT ACCGAGTATT
 ..........................................................................

6701  CACCCCTTGT ATTACTGTTT ATGTAAGCAG ACAGTTTTAT TGTTCATGAT
      GTGGGGAACA TAATGACAAA TACATTCGTC TGTCAAAATA ACAAGTACTA
 ..........................................................................

                                                      DraIII
                                                      ------
6751  GATATATTTT TATCTTGTGC AATGTAACAT CAGAGATTTT GAGACACAAC
      CTATATAAAA ATAGAACACG TTACATTGTA GTCTCTAAAA CTCTGTGTTG
 ..........................................................................
```

DraIII

```
6801  GTGGCTTTCC CCCCCCCCCC ATTATTGAAG CATTTATCAG GGTTATTGTC
      CACCGAAAGG GGGGGGGGGG TAATAACTTC GTAAATAGTC CCAATAACAG
 ..........................................................................
6851  TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAAATAGGG
      AGTACTCGCC TATGTATAAA CTTACATAAA TCTTTTTATT TGTTTATCCC
 ..........................................................................
6901  GTTCCGCCCA CATTTCCCCG AAAAGTGCCA CCTGACGTCT AAGAAACCAT
      CAAGGCGCGT GTAAAGGGGC TTTTCACGGT GGACTGCAGA TTCTTTGGTA
 ..........................................................................
6951  TACTATCATG ACATTAACCT ATAAAAATAG GCGTATCACG AGGCCCTTTC
      ATAATAGTAC TGTAATTGGA TATTTTTATC CGCATAGTGC TCCGGGAAAG
 ..........................................................................
7001  GTC
      CAG
 ..........................................................................
```

pVR 1012-GP(S)

General Description

DNA pVR 1012-GP(S)
Local object
Created: 09/14/98 03:58PM
Last Modification Date: ? (no data)
length: 7073 bp
storage type: Basic
form: Circular
Comments

Restriction Map

BalI: 1 site TGGCCA / ACCGGT

BclI: 1 site TGATCA / ACTAGT

ClaI: 1 site ATCGAT / TAGCTA

DraIII: 1 site CACNNNGTG / GTGNNNCAC

HindIII: 1 site AAGCTT / TTCGAA

KasI: 1 site GGCGCC / CCGCGG

KpnI: 1 site GGTACC / CCATGG

NarI: 1 site GGCGCC / CCGCGG

PmlI: 1 site CACGTG / GTGCAC

PvuI: 1 site CGATCG / GCTAGC

SacII: 1 site CCGCGG / GGCGCC

SalI: 1 site GTCGAC / CAGCTG

XbaI: 1 site TCTAGA / AGATCT

XmnI: 1 site GAANNNNTTC / CTTNNNNAAG

NdeI: 2 sites CATATG / GTATAC

EcoRV: 3 sites GATATC / CTATAG

SphI: 3 sites GCATGC / CGTACG

NcoI: 4 sites CCATGG / GGTACC

BamHI: 6 sites GGATCC / CCTAGG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4090 End: 4642

Kan r

Start: 6138 End: 6760 (Complementary)

Misc_feature (2 signals)

CMV enhancer

Start: 248 End: 885

GP(S)

Start: 1870 End: 4089

Annotations

```
   1  TCCCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

  51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
      CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

 101  TCAGGGCCCG TCAGCGGGTG TTGCCGGGTG TCGGGGCTGG CTTAACTATG
      AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                          NdeI
                                         ------

 151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
      GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACGC CACACTTTAT

                                                    BalI
                                                   ------

 201  CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
      GGCGTGTCTA CGCATTCCTC TTTTATGGCG TAGTCTAACC GATAACCGGT

 251  TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
      AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

 301  TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
      AGGTTGTAAT GGCGGTACAA CTGTAACTAA TAACTGATCA ATAATTATCA

 351  AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
      TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA

 401  ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG
      TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC

 451  CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
      GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTGC GGTTATCCCT

 501  CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
      GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTG ACGGGTGAAC

                        NdeI
                       ------

 551  GCAGTACATC AATGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
      CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT

 601  TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
      ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGGAATACCC

                                                    NcoI
                                                   ------

 651  ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
      TGAAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

      NcoI
      ----

 701  GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
      CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG

 751  ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
      TGCCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAAA
```

```
 801  GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
      CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT
 851  TTGACGCAAA TGGGCGGTAG GCGTGTACGG TGGGAGGTCT ATATAAGCAG
      AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC
 901  AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
      TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA
                                                 SacII
                                               ========
 951  TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
      AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT
1001  CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
      GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG
                                                 SphI
                                               ========
1051  CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
      GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA
1101  TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
      AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC
1151  TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
      ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG
1201  TATTCGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
      ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT
1251  CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
      GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG
1301  ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
      TGCCTGAGAC ATAAAAATGT CCTACCCCAG GGTAAATAAT AAATGTTTAA
1351  CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
      GTGTATATGT TGTTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT
1401  GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
      CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA
1451  TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
      AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG
1501  AGCGGCTCAT GGTCGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
      TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC
1551  ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
      TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC
1601  TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
      ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGAGCGTGC
1651  GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
      CGACTGCGTC TACCTTCTGA ATTCCGTCGC CGTCTTCTTC TACGTCCGTC
```

```
1701  CTGAGTTCTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCCGTGC
      GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG

1751  TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
      ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

                                                      NcoI
                                                      ----
1801  CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
      GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA

                              SalI
                              ----
      NcoI                        PmlI   BclI   EcoRV
      ----                        ----------------
1851  GCGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
      CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

                      SphI                            EcoRV
                      ----                            -----
1901  CCGCTCTAGC TAGATGCATG CTCGAGCGGC CGCCAGTGTG ATGGATATCT
      GGCGAGATCG ATCTACGTAC GAGCTCGCCG GCGGTCACAC TACCTATAGA

                              NcoI
                              ----
2951  GCAGAATTCT ATCTTCAGGA TCTCGCCATG GAGGGTCTTA GCCTACTCCA
      CGTCTTAAGA TAGAAGTCCT AGAGCGGTAC CTCCCAGAAT CGGATGAGGT

2001  ATTGCCCAGA GATAAATTTC GAAAAAGCTC TTTCTTTGTT TGGGTCATCA
      TAACGGGTCT CTATTTAAAG CTTTTTCGAG AAAGAAACAA ACCCAGTAGT

2051  TCTTATTTCA AAAGGCCTTT TCCATGCCTT TGGGTGTTGT GACCAACAGC
      AGAATAAAGT TTTCCGGAAA AGGTACGGAA ACCCACAACA CTGGTTGTCG

2101  ACTTTAGAAG TAACAGAGAT TGACCAGCTA GTCTGCAAGG ATCATCTTGC
      TGAAATCTTC ATTGTCTCTA ACTGGTCGAT CAGACGTTCC TAGTAGAACG

2151  ATCAACTGAC CAGCTGAAAT CAGTTGGTCT CAACCTCGAG GGGAGCGGAG
      TAGTTGACTG GTCGACTTTA GTCAACCAGA GTTGGAGCTC CCCTCGCCTC

              EcoRV
              -----
2201  TATCTACTGA TATCCCATCT GCGACAAAGC GTTGGGGCTT CAGATCTGGT
      ATAGATGACT ATAGGGTAGA CGCTGTTTCG CAACCCCGAA GTCTAGACCA

2251  GTGCCTCCCC AAGTGGTCAG CTATGAAGCA GGAGAATGGG CTGAAAATTG
      CACGGAGGGG TTCACCAGTC GATACTTCGT CCTCTTACCC GACTTTTAAC

2301  CTACAATCTT GAAATAAAGA AACCGGACGG GAGCGAATGC TTACCCCCAC
      GATGTTAGAA CTTTATTTCT TTGGCCTGCC CTCGCTTACG AATGGGGGTG

2351  CGCCGGATGG TGTCAGAGGC TTTCCAAGGT GCCGCTATGT TCACAAAGCC
      GCGGCCTACC ACAGTCTCCG AAAGGTTCCA CGGCGATACA AGTGTTTCGG

2401  CAAGGAACCG GGCCCTGCCC GGGTGACTAC GCCTTTCACA AGGATGGAGC
      GTTCCTTGGC CCGGGACGGG CCCACTGATA CGGAAAGTGT TCCTACCTCG

2451  TTTCTTCCTC CATGACAGGC TGGCTTCAAC TGTAATTTAC AGAGGAGTCA
      AAAGAAGGAG ATACTGTCCG ACCGAAGTTG ACATTAAATG TCTCCTCAGT
```

```
2501 ATTTTGCTGA GCGGGTAATC GCATTCTTGA TATTGGCTAA ACCAAAGGAA
     TAAAACGACT CGCCCATTAG CGTAAGAACT ATAACCGATT TGGTTTCCTT

2551 ACGTTCCTTC AATCACCCCC CATTCGAGAG GCAGCAAACT ACACTGAAAA
     TGCAAGGAAG TTAGTGGGGG GTAAGCTCTC CGTCGTTTGA TGTGACTTTT

2601 TACATCAAGT TACTATGCCA CATCCTACTT GGAGTACGAA ATCGAAAATT
     ATGTAGTTCA ATGATACGGT GTAGGATGAA CCTCATGCTT TAGCTTTTAA

2651 TTGGTGCTCA ACACTCCACG ACCCTTTTCA AAATTAACAA TAATACTTTT
     AACCACGAGT TGTGAGGTGC TGGGAAAAGT TTTAATTGTT ATTATGAAAA

2701 GTTCTTCTGG ACAGGCCCCA CACGCCTCAG TTCCTTTTCC AGCTGAATGA
     CAAGAAGACC TGTCCGGGGT GTGCGGAGTC AAGGAAAAGG TCGACTTACT

2751 TACCATTCAA CTTCACCAAC AGTTGAGCAA CACAACTGGG AAACTAATTT
     ATGGTAAGTT GAAGTGGTTG TCAACTCGTT GTGTTGACCC TTTGATTAAA

2801 GGACACTAGA TGCTAATATC AATGCTGATA TTGGTGAATG GGCTTTTTGG
     CCTGTGATCT ACGATTATAG TTACGACTAT AACCACTTAC CCGAAAAACC

2851 GAAAATAAAA AAATCTCTCC GAACAACTAC GTGGAGAAGA GCTGTCTTTC
     CTTTTATTTT TTTAGAGAGG CTTGTTGATG CACCTCTTCT CGACAGAAAG

2901 GAAACTTTAT CGCTCAACGA GACAGAAGAC GATGATGCGA CATCGTCGAG
     CTTTGAAATA GCGAGTTGCT CTGTCTTCTG CTACTACGCT GTAGCAGCTC

2951 AACTACAAAG GGAAGAATCT CCGACCGGGC CACCAGGAAG TATTCGGACC
     TTGATGTTTC CCTTCTTAGA GGCTGGCCCG GTGGTCCTTC ATAAGCCTGG

3001 TGGTTCCAAA GGATTCCCCT GGGATGGTTT CATTGCACGT ACCAGAAGGG
     ACCAAGGTTT CCTAAGGGGA CCCTACCAAA GTAACGTGCA TGGTCTTCCC

3051 GAAACAACAT TGCCGTCTCA GAATTCGACA GAAGGTCGAA GAGTAGATGT
     CTTTGTTGTA ACGGCAGAGT CTTAAGCTGT CTTCCAGCTT CTCATCTACA

3101 GAATACTCAG GAAACTATCA CAGAGACAAC TGCAACAATC ATAGGCACTA
     CTTATGAGTC CTTTGATAGT GTCTCTGTTG ACGTTGTTAG TATCCGTGAT

3151 ACGGTAACAA CATGCAGATC TCCACCATCG GGACAGGACT GAGCTCCAGC
     TGCCATTGTT GTACGTCTAG AGGTGGTAGC CCTGTCCTGA CTCGAGGTCG

                                NcoI
3201 CAAATCCTGA GTTCCTCACC GACCATGGCA CCAAGCCCTG AGACTCAGAC
     GTTTAGGACT CAAGGAGTGG CTGGTACCGT GGTTCGGGAC TCTGAGTCTG

3251 CTCCACAACC TACACACCAA AACTACCAGT GATGACCACC GAGGAACCAA
     GAGGTGTTGG ATGTGTGGTT TTGATGGTCA CTACTGGTGG CTCCTTGGTT

3301 CAACACCACC GAGAAACTCT CCTGGCTCAA CAACAGAAGC ACCCACTCTC
     GTTGTGGTGG CTCTTTGAGA GGACCGAGTT GTTGTCTTCG TGGGTGAGAG

3351 ACCACCCCAG AGAATATAAC AACACCGGTT AAAACTGTTT GGGCACAAGA
     TGGTGGGGTC TCTTATATTG TTGTGGCCAA TTTTGACAAA CCCGTGTTCT
```

```
3401 GTCCACAAGC AACGGTCTAA TAACTTCAAC AGTAACAGGT ATTCTTGGGA
     CAGGTGTTCG TTGCCAGATT ATTGAAGTTG TCATTGTCCA TAAGAACCCT

3451 GCCTTGGACT TCGAAAACGC AGCAGAAGAC AAGTTAACAC CAGGGCCACG
     CGGAACCTGA AGCTTTTGCG TCGTCTTCTG TTCAATTGTG GTCCCGGTGC

3501 GGTAAATGCA ATCCCAACTT ACACTACTGG ACTGCACAAG AACAACATAA
     CCATTTACGT TAGGGTTGAA TGTGATGACC TGACGTGTTC TTGTTGTATT

                   BamHI
3551 TGCTGCTGGG ATTGCCTGGA TCCCGTACTT TGGACCGGGT GCAGAAGGCA
     ACGACGACCC TAACGGACCT AGGGCATGAA ACCTGGCCCA CGTCTTCCGT

3601 TATACACTGA AGGCCTTATG CACAACCAAA ATGCCTTAGT CTGTGGACTC
     ATATGTGACT TCCGGAATAC GTGTTGGTTT TACGGAATCA GACACCTGAG

3651 AGACAACTTG CAAATGAAAC AACTCAAGCT CTGCAGCTTT TCTTAAGGGC
     TCTGTTGAAC GTTTACTTTG TTGAGTTCGA GACGTCGAAA AGAATTCCCG

3701 CACGACGGAG CTGCGGACAT ATACCATACT CAATAGGAAG GCCATAGATT
     GTGCTGCCTC GACGCCTGTA TATGGTATGA GTTATCCTTC CGGTATCTAA

                                       BamHI
3751 TCCTTCTGCG ACGATGGGCC GGGACATGTA GGATCCTGGG ACCAGATTGT
     AGGAAGACGC TGCTACCCCG CCCTGTACAT CCTAGGACCC TGGTCTAACA

3801 TGCATTGAGC CACATGATTG GACCAAAAAC ATCACTGATA AAATCAACCA
     ACGTAACTCG GTGTACTAAC CTGGTTTTTG TAGTGACTAT TTTAGTTGGT

3851 AATCATCCAT GATTTCATCG ACAACCCTTT ACCCAATCAG GATAATGATG
     TTAGTAGGTA CTAAAGTAGC TGTTGGGAAA TGGGTTAGTC CTATTACTAC

                                  BamHI
3901 ATAATTGGTG GACGGGCTGG AGACAGTGGA TCCCTGCAGG AATAGGCATT
     TATTAACCAC CTGCCCGACC TCTGTCACCT AGGGACGTCC TTATCCGTAA

3951 ACTGGAATTA TTATTGCAAT CATTGCTCTT CTTTGCGTCT GCAAGCTGCT
     TGACCTTAAT AATAACGTTA GTAACGAGAA GAAACGCAGA CGTTCGACGA

                                                  BamHI
4001 TTGTTGAATA TCAGAATTCC AGCACTGGCG GCCGTTACTA GTGGATCCGA
     AACAACTTAT AGTCTTAAGG TCGTGACCGC CGGCAATGAT CACCTAGGCT

                            NarI
        BamHI     XbaI      KasI   BamHI
4051 GCTCGGATCC AAGCTCTAGA CCAGGCGCCT GGATCCAGAT CTGCTGTGCC
     CGAGCCTAGG TTCGAGATCT GGTCCGCGGA CCTAGGTCTA GACGACACGG

4101 TTCTAGTTGC CAGCCATCTG TTGTTTGCCC CTCCCCCGTG CCTTCCTTGA
     AAGATCAACG GTCGGTAGAC AACAAACGGG GAGGGGGCAC GGAAGGAACT

4151 CCCTGGAAGG TGCCACTCCC ACTGTCCTTT CCTAATAAAA TGAGGAAATT
     GGGACCTTCC ACGGTGAGGG TGACAGGAAA GGATTATTTT ACTCCTTTAA
```

```
4201  GCATCGCATT GTCTGAGTAG GTGTCATTCT ATTCTGGGGG GTGGGGTGGG
      CGTAGCGTAA CAGACTCATC CACAGTAAGA TAAGACCCCC CACCCCACCC

                                                    SphI
                                                   ------
4251  GCAGCACAGC AAGGCGGAGG ATTGGGAAGA CAATAGCAGG CATGCTGGGG
      CGTCGTGTCG TTCCCCCTCC TAACCCTTCT GTTATCGTCC GTACGACCCC

                      KpnI
                     ------
4301  ATGCGGTGGG CTCTATGGGT ACCCAGGTGC TGAAGAATTC ACCCGGTTCC
      TACGCCACCC GAGATACCCA TGGGTCCACG ACTTCTTAAG TGGGCCAAGG

4351  TCCTGGGCCA GAAAGAAGCA GGCACATCCC CTTCTCTGTG ACACACCCTG
      AGGACCCGGT CTTTCTTCGT CCGTGTAGGG GAAGAGACAC TGTGTGGGAC

4401  TCCACGCCCC TGGTTCTTAG TTCCAGCCCC ACTCATAGGA CACTCATAGC
      AGGTGCGGGG ACCAAGAATC AAGGTCGGGG TGAGTATCCT GTGAGTATCG

4451  TCAGGAGGGC TCCGCCTTCA ATCCCACCCG CTAAAGTACT TGGAGCGGTC
      AGTCCTCCCG AGGCGGAAGT TAGGGTGGGC GATTTCATGA ACCTCGCCAG

4501  TCTCCCTCCC TCATCAGCCC ACCAAACCAA ACCTAGCCTC CAAGAGTGGG
      AGAGGGAGGG AGTAGTCGGG TGGTTTGGTT TGGATCGGAG GTTCTCACCC

4551  AAGAAATTAA AGCAAGATAG GCTATTAAGT GCAGAGGGAG AGAAAATGCC
      TTCTTTAATT TCGTTCTATC CGATAATTCA CGTCTCCCTC TCTTTTACGG

                                          XmnI
                                         ------
4601  TCCAACATGT GAGGAAGTAA TCAGAGAAAT CATAGAATTT CTTCCGCCTC
      AGGTTGTACA CTCCTTCATT AGTCTCTTTA GTATCTTAAA GAAGGCGAAG

4651  CTCGCTCACT GACTCGCTGC GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT
      GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA

4701  CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC
      GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG

4751  GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA
      CGTCCTTTCT TGTACACTCG TTTTCCGGTC GTTTTCCGGT CCTTGGCATT

4801  AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC
      TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGACTGCTCG

4851  ATCACAAAAA TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA
      TAGTGTTTTT AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCCTGAT

4901  TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT
      ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCACG CGAGAGGACA

4951  TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA
      AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT

5001  GCGTGGCGCT TTCTCAATGC TCACGCTGTA GGTATCTCAG TTCGGTGTAG
      CGCACCGCGA AAGAGTTACG AGTGCGACAT CCATAGAGTC AAGCCACATC
```

```
5051  GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA
      CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT
5101  CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC
      GGCGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG
5151  ACGACTTATC GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG
      TGCTGAATAG CGGTGACCGT CGTCGGTGAC CATTGTCCTA ATCGTCTCGC
5201  AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACTACGG
      TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC
5251  CTACACCAGA AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA
      GATGTGATCT TCCTGTCATA AACCATAGAC GCGAGACGAC TTCGGTCAAT
5301  CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA AACCACCGCT
      GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT TTGGTGGCGA
5351  GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA
      CCATCGCCAC CAAAAAAACA AACGTTCGTC GTCTAATGCG CGTCTTTTTT
5401  AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT
      TCCTAGAGTT CTTCTAGGAA ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA
5451  GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG
      CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC
5501  ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA
      TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTCAGAT
5551  AAGTATATAT GAGTAAACTT GGTCTGACAG TTACCAATGC TTAATCAGTG
      TTCATATATA CTCATTTGAA CCAGACTGTC AATGGTTACG AATTAGTCAC
5601  AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCTGA
      TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT
5651  CTCCGGGGGG GGGGGGCGCT GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG
      GAGGCCCCCC CCCCCCGCGA CTCCAGACGG AGCACTTCTT CCACAACGAC
5701  ACTCATACCA GGCCTGAATC GCCCCATCAT CCAGCCAGAA AGTGAGGGAG
      TGAGTATGGT CCGGACTTAG CGGGGTAGTA GGTCGGTCTT TCACTCCCTC
5751  CCACGGTTGA TGAGAGCTTT GTTGTAGGTG GACCAGTTGG TGATTTTGAA
      GGTGCCAACT ACTCTCGAAA CAACATCCAC CTGGTCAACC ACTAAAACTT
5801  CTTTTGCTTT GCCACGGAAC GGTCTGCGTT GTCGGGAAGA TGCGTGATCT
      GAAAACGAAA CGGTGCCTTG CCAGACGCAA CAGCCCTTCT ACGCACTAGA
5851  GATCCTTCAA CTCAGCAAAA GTTCGATTTA TTCAACAAAG CCGCCGTCCC
      CTAGGAAGTT GAGTCGTTTT CAAGCTAAAT AAGTTGTTTC GGCGGCAGGG
5901  GTCAAGTCAG CGTAATGCTC TGCCAGTGTT ACAACCAATT AACCAATTCT
      CAGTTCAGTC GCATTACGAG ACGGTCACAA TGTTGGTTAA TTGGTTAAGA
5951  GATTAGAAAA ACTCATCGAG CATCAAATGA AACTGCAATT TATTCATATC
      CTAATCTTTT TGAGTAGCTC GTAGTTTACT TTGACGTTAA ATAAGTATAG
```

```
6001  AGGATTATCA ATACCATATT TTTGAAAAAG CCGTTTCTGT AATGAAGGAG
      TCCTAATAGT TATGGTATAA AAACTTTTTC GGCAAAGACA TTACTTCCTC

6051  AAAACTCACC GAGGCAGTTC CATACGATGG CAAGATCCTG GTATCGGTCT
      TTTTGAGTGG CTCCGTCAAG GTATCCTACC GTTCTAGGAC CATAGCCAGA

6101  GCGATTCCGA CTCGTCCAAC ATCAATACAA CCTATTAATT TCCCCTCGTC
      CGCTAAGGCT GAGCAGGTTG TAGTTATGTT GGATAATTAA AGGGGAGCAG

6151  AAAAATAAGG TTATCAAGTG AGAAATCACC ATGAGTGACG ACTGAATCCG
      TTTTTATTCC AATAGTTCAC TCTTTAGTGG TACTCACTGC TGACTTAGGC

                  HindIII
                  ========
6201  GTGAGAATGG CAAAAGCTTA TGCATTTCTT TCCAGACTTG TTCAACAGGC
      CACTCTTACC GTTTTCGAAT ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG

6251  CAGCCATTAC GCTCGTCATC AAAATCACTC GCATCAACCA AACCGTTATT
      GTCGGTAATG CGAGCAGTAG TTTTAGTGAG CGTAGTTGGT TTGGCAATAA

                                        PvuI
                                        ====
6301  CATTCGTGAT TCCGCCTGAG CGAGACGAAA TACGCGATCG CTGTTAAAAG
      GTAAGCACTA ACGCGGACTC GCTCTGCTTT ATGCGCTAGC GACAATTTTC

6351  GACAATTACA AACAGGAATC GAATGCAACC GGCGCAGGAA CACTGCCAGC
      CTGTTAATGT TTGTCCTTAG CTTACGTTGG CCGCGTCCTT GTGACGGTCG

6401  GCATCAACAA TATTTTCACC TGAATCAGGA TATTCTTCTA ATACCTGGAA
      CGTAGTTGTT ATAAAAGTGG ACTTAGTCCT ATAAGAAGAT TATGGACCTT

6451  TGCTGTTTTC CCGGGGATCG CAGTGGTGAG TAACCATGCA TCATCAGGAG
      ACGACAAAAG GGCCCCTAGC GTCACCACTC ATTGGTACGT AGTAGTCCTC

6501  TACGGATAAA ATGCTTGATG GTCGGAAGAG GCATAAATTC CGTCAGCCAG
      ATGCCTATTT TACGAACTAC CAGCCTTCTC CGTATTTAAG GCAGTCGGTC

6551  TTTAGTCTGA CCATCTCATC TGTAACATCA TTGGCAACGC TACCTTTGCC
      AAATCAGACT GGTAGAGTAG ACATTGTAGT AACCGTTGCG ATGGAAACGG

                                                    ClaI
                                                    ====
6601  ATGTTTCAGA AACAACTCTG GCGCATCGGG CTTCCCATAC AATCGATAGA
      TACAAAGTCT TTGTTGAGAC CGCGTAGCCC GAAGGGTATG TTAGCTATCT

6651  TTGTCGCACC TGATTGCCCG ACATTATCGC GAGCCCATTT ATACCCATAT
      AACAGCGTGG ACTAACGGGC TGTAATAGCG CTCGGGTAAA TATGGGTATA

6701  AAATCAGCAT CCATGTTGGA ATTTAATCGC GGCCTCGAGC AAGACGTTTC
      TTTAGTCGTA GGTACAACCT TAAATTAGCG CCGGAGCTCG TTCTGCAAAG

6751  CCGTTGAATA TGGCTCATAA CACCCCTTGT ATTACTGTTT ATGTAAGCAG
      GGCAACTTAT ACCGAGTATT GTGGGGAACA TAATGACAAA TACATTCGTC

6801  ACAGTTTTAT TGTTCATGAT GATATATTTT TATCTTGTGC AATGTAACAT
      TGTCAAAATA ACAAGTACTA CTATATAAAA ATAGAACACG TTACATTGTA
```

DraIII

```
6851  CAGAGATTTT GAGACACAAC GTGGCTTTCC CCCCCCCCCC ATTATTGAAG
      GTCTCTAAAA CTCTGTGTTG CACCGAAAGG GGGGGGGGGG TAATAACTTC
 ....................................................................................
6901  CATTTATCAG GGTTATTGTC TCATGAGCGG ATACATATTT GAATGTATTT
      GTAAATAGTC CCAATAACAG AGTACTCGCC TATGTATAAA CTTACATAAA
 ....................................................................................
6951  AGAAAAATAA ACAAATAGGG GTTCCGCGCA CATTTCCCCG AAAAGTGCCA
      TCTTTTTATT TGTTTATCCC CAAGGCGCGT GTAAAGGGGC TTTTCACGGT
 ....................................................................................
7001  CCTGACGTCT AAGAAACCAT TATTATCATG ACATTAACCT ATAAAAATAG
      GGACTGCAGA TTCTTTGGTA ATAATAGTAC TGTAATTGGA TATTTTTATC
 ....................................................................................
7051  GCGTATCACG AGGCCCTTTC GTC
      CGCATAGTGC TCCGGGAAAG CAG
 ....................................................................................
```

SEQUENCE LISTING ID NO: 3

pVR 1012-GP(Z)

General Description

DNA pVR 1012-GP(Z)
Local object
Created: 09/15/98 05:06PM
Last Modification Date: ? (no data)
length: 7285 bp
storage type: Basic
form.: Circular
Comments

Restriction Map

DraIII: 1 site CACNNNGTG / GTGNNNCAC

HindIII: 1 site AAGCTT / TTCGAA

HpaI: 1 site GTTAAC / CAATTG

KasI: 1 site GGCGCC / CCGCGG

NarI: 1 site GGCGCC / CCGCGG

NotI: 1 site GCGGCCGC / CGCCGGCG

PmlI: 1 site CACGTG / GTGCAC

PvuI: 1 site CGATCG / GCTAGC

SacII: 1 site CCGCGG / GGCGCC

XbaI: 1 site TCTAGA / AGATCT

XhoI: 1 site CTCGAG / GAGCTC

EcoRV: 2 sites GATATC / CTATAG

NcoI: 2 sites CCATGG / GGTACC

NdeI: 2 sites CATATG / GTATAC

SphI: 2 sites GCATGC / CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT
Start: 886 End: 1129

CMV IE INT
Start: 1130 End: 1860

TbGH
Start: 4302 End: 4854

Kan r
Start: 6350 End: 6972 (Complementary)

Misc_feature (2 signals)

CMV enhancer

Start: 248 End: 885

GP(Z)

Start: 1870 End: 4301

Annotations

```
   1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

  51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
      CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

 101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
      AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                          NdeI
                                         ------
 151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
      GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACGC CACACTTTAT

 201  CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
      GGCGTGTCTA CGCATTCCTC TTTTATGGCG TAGTCTAACC GATAACCGGT

 251  TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
      AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

 301  TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
      AGGTTGTAAT GGCGGTACAA CTGTAACTAA TAACTGATCA ATAATTATCA

 351  AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
      TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA

 401  ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG
      TGTATTGAAT CCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC

 451  CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
      GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTGC GGTTATCCCT

 501  CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
      GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTG ACGGGTGAAC

                         NdeI
                        ------
 551  GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
      CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT

 601  TGACGGTAAA TGGCCCCCCT GGCATTATGC CCAGTACATG ACCTTATGGG
      ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGGAATACCC

                                                    NcoI
                                                   ------
 651  ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
      TGAAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

      NcoI
      -
 701  GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
      CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG

 751  ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
      TGCCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAAA

 801  GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
      CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT
```

```
 851  TTGACGCAAA TGGCCGGTAG GCGTGTACGG TGGGAGGTCT ATATAAGCAG
      AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC
 ..........................................................................

 901  AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
      TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA
 ..........................................................................

                                                 SacII
                                                 ------

 951  TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
      AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT
 ..........................................................................

1001  CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
      GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG
 ..........................................................................

                                                 SphI
                                                 ----

1051  CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
      GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA
 ..........................................................................

1101  TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
      AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC
 ..........................................................................

1151  TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
      ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG
 ..........................................................................

1201  TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
      ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT
 ..........................................................................

1251  CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
      GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG
 ..........................................................................

1301  ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
      TGCCTGAGAC ATAAAAATGT CCTACCCCAG GGTAAATAAT AAATGTTTAA
 ..........................................................................

1351  CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
      GTGTATATGT TGTTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT
 ..........................................................................

1401  GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
      CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA
 ..........................................................................

1451  TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
      AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG
 ..........................................................................

1501  AGCGGCTCAT GGTCGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
      TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC
 ..........................................................................

1551  ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
      TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC
 ..........................................................................

1601  TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
      ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGAGCGTGC
 ..........................................................................

1651  GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
      CGACTGCGTC TACCTTCTGA ATTCCGTCGC CGTCTTCTTC TACGTCCGTC
 ..........................................................................

1701  CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
      GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG
 ..........................................................................
```

HpaI

1751 TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

NcoI

1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA

NcoI PmlI EcoRVNotI

1851 GGGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

NarI

NotI XbaI KasI

1901 CCGCTCTAGA CCAGGCGCCT GGATCGATCC GCGATGAAGA TTAAGCCGAC
GGCGAGATCT GGTCCGCGGA CCTAGCTAGG CGCTACTTCT AATTCGGCTG

1951 AGTGAGCGTA ATCTTCATCT CTCTTAGATT ATTTGTTTTC CAGAGTAGGG
TCACTCGCAT TAGAAGTAGA GAGAATCTAA TAAACAAAAG GTCTCATCCC

2001 GTCGTCAGGT CCTTTTCAAT CGTGTAACCA AAATAAACTC CACTAGAAGG
CAGCAGTCCA GGAAAAGTTA GCACATTGGT TTTATTTGAG GTGATCTTCC

2051 ATATTGTGGG GCAACAACAC AATGGGCGTT ACAGGAATAT TCCAGTTACC
TATAACACCC CGTTGTTGTG TTACCCGCAA TGTCCTTATA AGGTCAATGG

2101 TCGTGATCGA TTCAAGAGGA CATCATTCTT TCTTTGGGTA ATTATCCTTT
AGCACTAGCT AAGTTCTCCT GTAGTAAGAA AGAAACCCAT TAATAGGAAA

2151 TCCAAAGAAC ATTTTCCATC CCACTTGGAG TCATCCACAA TAGCACATTA
AGGTTTCTTG TAAAAGGTAG GGTGAACCTC AGTAGGTGTT ATCGTGTAAT

2201 CAGGTTAGTG ATGTCGACAA ACTAGTTTGT CGTGACAAAC TGTCATCCAC
GTCCAATCAC TACAGCTGTT TGATCAAACA GCACTGTTTG ACAGTAGGTG

2251 AAATCAATTG AGATCAGTTG GACTGAATCT CGAAGGGAAT GGAGTGGCAA
TTTAGTTAAC TCTAGTCAAC CTGACTTAGA GCTTCCCTTA CCTCACCGTT

2301 CTGACGTGCC ATCTGCAACT AAAAGATGGG GCTTCAGGTC CGGTGTCCCA
GACTGCACGG TAGACGTTGA TTTTCTACCC CGAAGTCCAG GCCACAGGGT

2351 CCAAAGGTGG TCAATTATGA AGCTGGTGAA TGGGCTGAAA ACTGCTACAA
GGTTTCCACC AGTTAATACT TCGACCACTT ACCCGACTTT TGACGATGTT

2401 TCTTGAAATC AAAAAACCTG ACGGGAGTGA GTGTCTACCA GCAGCGCCAG
AGAACTTTAG TTTTTTGGAC TGCCCTCACT CACAGATGGT CGTCGCGGTC

2451 ACGGGATTCG GGGCTTCCCC CGGTGCCGGT ATGTCCACAA AGTATCAGGA
TGCCCTAAGC CCCGAAGGGG GCCACGGCCA TACACGTGTT TCATAGTCCT

2501 ACGGGACCGT GTGCCGGAGA CTTTGCCTTC CATAAAGAGG GTGCTTTCTT
TGCCCTGGCA CACGGCCTCT GAAACGGAAG GTATTTCTCC CACGAAAGAA

```
2551  CCTGTATGAT CGACTTGCTT CCACAGTTAT CTACCGAGGA ACGACTTTCG
      GGACATACTA GCTGAACGAA GGTGTCAATA GATGGCTCCT TGCTGAAAGC

2601  CTGAAGGTGT CGTTGCATTT CTGATACTGC CCCAAGCTAA GAAGGACTTC
      GACTTCCACA GCAACGTAAA GACTATGACG GGGTTCGATT CTTCCTGAAG

2651  TTCAGCTCAC ACCCCTTGAG AGAGCCGGTC AATGCAACGG AGGACCCGTC
      AAGTCGAGTG TGGGGAACTC TCTCGGCCAG TTACGTTGCC TCCTGGGCAG

                                  EcoRV
                                 ------
2701  TAGTGGCTAC TATTCTACCA CAATTAGATA TCAGGCTACC GGTTTTGGAA
      ATCACCGATG ATAAGATGGT GTTAATCTAT AGTCCGATGG CCAAAACCTT

2751  CCAATGAGAC AGAGTACTTG TTCGAGGTTG ACAATTTGAC CTACGTCCAA
      GGTTACTCTG TCTCATGAAC AAGCTCCAAC TGTTAAACTG GATGCAGGTT

2801  CTTGAATCAA GATTCACACC ACAGTTTCTG CTCCAGCTGA ATGAGACAAT
      GAACTTAGTT CTAAGTGTGG TGTCAAAGAC GAGGTCGACT TACTCTGTTA

2851  ATATACAAGT GGGAAAAGGA GCAATACCAC GGGAAAACTA ATTTGGAAGG
      TATATGTTCA CCCTTTTCCT CGTTATGGTG CCCTTTTGAT TAAACCTTCC

2901  TCAACCCCGA AATTGATACA ACAATCGGGG AGTGGGCCTT CTGGGAAACT
      AGTTGGGGCT TTAACTATGT TGTTAGCCCC TCACCCGGAA GACCCTTTGA

2951  AAAAAAAACC TCACTAGAAA AATTCGCAGT GAAGAGTTGT CTTTCACAGT
      TTTTTTTTGG AGTGATCTTT TTAAGCGTCA CTTCTCAACA GAAAGTGTCA

3001  TGTATCAAAC GGAGCCAAAA ACATCAGTGG TCAGAGTCCG GCGCGAACTT
      ACATAGTTTG CCTCGGTTTT TGTAGTCACC AGTCTCAGGC CGCGCTTGAA

3051  CTTCCGACCC AGGGACCAAC ACAACAACTG AAGACCACAA AATCATGGCT
      GAAGGCTGGG TCCCTGGTTG TGTTGTTGAC TTCTGGTGTT TTAGTACCGA

3101  TCAGAAAATT CCTCTGCAAT GGTTCAAGTG CACAGTCAAG GAAGGGAAGC
      AGTCTTTTAA GGAGACGTTA CCAAGTTCAC GTGTCAGTTC CTTCCCTTCG

3151  TGCAGTGTCG CATCTAACAA CCCTTGCCAC AATCTCCACG AGTCCCCAAT
      ACGTCACAGC GTAGATTGTT GGGAACGGTG TTAGAGGTGC TCAGGGGTTA

3201  CCCTCACAAC CAAACCAGGT CCGGACAACA GCACCCATAA TACACCCGTG
      GGGAGTGTTG GTTTGGTCCA GGCCTGTTGT CGTGGGTATT ATGTGGGCAC

3251  TATAAACTCG ACATCTCTGA GGCAACTCAA GTTGAACAAC ATCACCGCAG
      ATATTTGAAC TGTAGAGACT CCGTTGAGTT CAACTTGTTG TAGTGGCGTC

3301  AACAGACAAC GACAGCACAG CCTCCGACAC TCCCTCTGCC ACGACCGCAG
      TTGTCTGTTG CTGTCGTGTC GGAGGCTGTG AGGGAGACGG TGCTGGCGTC

3351  CCGGACCCCC AAAAGCAGAG AACACCAACA CGAGCAAGAG CACTGACTTC
      GGCCTGGGGG TTTTCGTCTC TTGTGGTTGT GCTCGTTCTC GTGACTGAAG

3401  CTGGACCCCG CCACCACAAC AAGTCCCCAA AACCACAGCG AGACCGCTGG
      GACCTGGGGC GGTGGTGTTG TTCAGGGGTT TTGGTGTCGC TCTGGCGACC
```

```
3451  CAACAACAAC ACTCATCACC AAGATACCGG AGAAGAGAGT GCCAGCAGCG
      GTTGTTGTTG TGAGTAGTGG TTCTATGGCC TCTTCTCTCA CGGTCGTCGC

3501  GGAAGCTAGG CTTAATTACC AATACTATTG CTGGAGTCGC ACGACTGATC
      CCTTCGATCC GAATTAATGG TTATGATAAC GACCTCAGCG TCCTGACTAG

3551  ACAGGCGGGA GAAGAACTCG AAGAGAAGCA ATTGTCAATG CTCAACCCAA
      TGTCCGCCCT CTTCTTGAGC TTCTCTTCGT TAACAGTTAC GAGTTGGGTT

3601  ATCCAACCCT AATTTACATT ACTGGACTAC TCAGGATGAA GGTGCTGCAA
      TAGGTTGGGA TTAAATGTAA TGACCTGATG AGTCCTACTT CCACGACGTT

3651  TCGGACTGGC CTGGATACCA TATTTCCGGC CAGCAGCCGA GGCAATTTAC
      AGCCTGACCG GACCTATGGT ATAAAGCCCG GTCGTCGGCT CCGTTAAATG

3701  ATAGAGGGGC TAATGCACAA TCAAGATGGT TTAATCTGTG GGTTGAGACA
      TATCTCCCCG ATTACGTGTT AGTTCTACCA AATTAGACAC CCAACTCTGT

3751  GCTGGCCAAC GAGACGACTC AAGCTCTTCA ACTGTTCCTG AGAGCCACAA
      CGACCGGTTG CTCTGCTGAG TTCGAGAAGT TGACAAGGAC TCTCGGTGTT

3801  CTGAGCTACG CACCTTTTCA ATCCTCAACC GTAAGGCAAT TGATTCCTTG
      GACTCGATGC GTGGAAAAGT TAGGAGTTGG CATTCCGTTA ACTAAAGAAC

3851  CTGCAGCGAT GGGGCGGCAC ATGCCACATT CTGGGACCGG ACTGCTGTAT
      GACGTCGCTA CCCCGCCGTG TACGGTGTAA GACCCTGGCC TGACGACATA

3901  CGAACCACAT GATTGGACCA AGAACATAAC AGACAAAATT GATCAGATTA
      GCTTGGTGTA CTAACCTGGT TCTTGTATTG TCTGTTTTAA CTAGTCTAAT

3951  TTCATGATTT TGTTGATAAA ACCCTTCCGG ACCAGGGGGA CAATGACAAT
      AAGTACTAAA ACAACTATTT TGGGAAGGCC TGGTCCCCCT GTTACTGTTA

4001  TGGTGGACAG GATGGAGACA ATGGATACCG GCAGGTATTG GAGTTACAGG
      ACCACCTGTC CTACCTCTGT TACCTATGGC CGTCCATAAC CTCAATGTCC

4051  CGTTATAATT GCAGTTATCG CTTTATTCTG TATATGCAAA TTTGTCTTTT
      GCAATATTAA CGTCAATAGC GAAATAAGAC ATATACGTTT AAACAGAAAA

4101  AGTTTTTCTT CAGATTGCTT CATGGAAAAG CTCAGCCTCA AATCAATGAA
      TCAAAAAGAA GTCTAACGAA GTACCTTTTC GAGTCGGAGT TTAGTTACTT

4151  ACCAGGATTT AATTATATGG ATTACTTGAA TCTAAGATTA CTTGACAAAT
      TGGTCCTAAA TTAATATACC TAATGAACTT AGATTCTAAT GAACTGTTTA

4201  GATAATATAA TACACTGGAG CTTTAAACAT AGCCAATGTG ATTCTAACTC
      CTATTATATT ATGTGACCTC GAAATTTGTA TCGGTTACAC TAAGATTGAG

4251  CTTTAAACTC ACAGTTAATC ATAAACAAGG TTTGGTACCG AGCTCGAATT
      GAAATTTGAG TGTCAATTAG TATTTGTTCC AAACCATGGC TCGAGCTTAA

4301  ATCTGCTGTG CCTTCTAGTT GCCAGCCATC TGTTGTTTGC CCCTCCCCCG
      TAGACGACAC GGAAGATCAA CGGTCGGTAG ACAACAAACG GGGAGGGGGC

4351  TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA
      ACGGAAGGAA CTGGGACCTT CCACGGTGAG GGTGACAGGA AAGGATTATT
```

```
4401  AATCAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG
      TTACTCCTTT AACGTAGCGT AACAGACTCA TCCACAGTAA GATAAGACCC
 .....

4451  GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATTGGGAA GACAATAGCA
      CCCACCCCAC CCCGTCGTGT CGTTCCCCCT CCTAACCCTT CTGTTATCGT
 .....

        SphI
       ------
4501  GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCCAGGT GCTGAAGAAT
      CCGTACGACC CCTACGCCAC CCGAGATACC CATGGGTCCA CGACTTCTTA
 .....

4551  TGACCCGGTT CCTCCTGGGC CAGAAAGAAG CAGGCACATC CCCTTCTCTG
      ACTGGGCCAA GGAGGACCCG GTCTTTCTTC GTCCGTGTAG GGGAAGAGAC
 .....

4601  TGACACACCC TGTCCACGCC CCTGGTTCTT AGTTCCAGCC CCACTCATAG
      ACTGTGTGGG ACAGGTGCGG GGACCAAGAA TCAAGGTCGG GGTGAGTATC
 .....

4651  GACACTCATA GCTCAGGAGG GCTCCGCCTC CAATCCCACC CGCTAAAGTA
      CTGTGAGTAT CGAGTCCTCC CGAGGCGGAA GTTAGGGTGG GCGATTTCAT
 .....

4701  CTTGGAGCGG TCTCTCCCTC CCTCATCAGC CCACCAAACC AAACCTAGCC
      GAACCTCGCC AGAGAGGGAG GGAGTAGTCG GGTGGTTTGG TTTGGATCGG
 .....

4751  TCCAAGAGTG GGAAGAAATT AAAGCAAGAT AGGCTATTAA GTGCAGAGGG
      AGGTTCTCAC CCTTCTTTAA TTTCGTTCTA TCCGATAATT CACGTCTCCC
 .....

4801  AGAGAAAATG CCTCCAACAT GTGAGGAAGT AATGAGAGAA ATCATAGAAT
      TCTCTTTTAC GGAGGTTGTA CACTCCTTCA TTACTCTCTT TAGTATCTTA
 .....

4851  TTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC
      AAGAAGGCGA AGGAGCGAGT GACTGAGCGA CGCGAGCCAG CAAGCCGACG
 .....

4901  GGCCAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA
      CCGCTCGCCA TAGTCGAGTG AGTTTCCGCC ATTATGCCAA TAGGTGTCTT
 .....

4951  TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC
      AGTCCCCTAT TGCGTCCTTT CTTGTACACT CGTTTTCCGG TCGTTTTCCG
 .....

5001  CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC
      GTCCTTGGCA TTTTTCCGGC GCAACGACCG CAAAAAGGTA TCCGAGGCGG
 .....

5051  CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC
      GGGGACTGCT CGTAGTGTTT TTAGCTGCGA GTTCAGTCTC CACCGCTTTG
 .....

5101  CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT
      GGCTGTCCTG ATATTTCTAT GGTCCGCAAA GGGGGACCTT CGAGGGAGCA
 .....

5151  GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC
      CGCGAGAGGA CAAGGCTGGG ACGGCGAATG GCCTATGGAC AGGCGGAAAG
 .....

5201  TCCCTTCGGG AAGCGTGGCG CTTTCTCAAT GCTCACGCTG TAGGTATCTC
      AGGGAAGCCC TTCGCACCGC GAAAGAGTTA CGAGTGCGAC ATCCATAGAG
 .....

5251  AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCC
      TCAAGCCACA TCCAGCAAGC GAGGTTCGAC CCGACACACG TGCTTGGGGG
 .....
```

```
5301  CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA
      GCAAGTCGGG CTGGCGACGC GGAATAGGCC ATTGATAGCA GAACTCAGGT
 .....
5351  ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG
      TCGGCCATTC TGTGCTGAAT AGCGGTGACC GTCGTCGGTG ACCATTGTCC
 .....
5401  ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG
      TAATCGTCTC GCTCCATACA TCCGCCACGA TGTCTCAAGA ACTTCACCAC
 .....
5451  GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC
      CGGATTGATG CCGATGTGAT CTTCCTGTCA TAAACCATAG ACGCGAGACG
 .....
5501  TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA
      ACTTCGGTCA ATGGAAGCCT TTTTCTCAAC CATCGAGAAC TAGGCCGTTT
 .....
5551  CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC
      GTTTGGTGGC GACCATCGCC ACCAAAAAAA CAAACGTTCG TCGTCTAATG
 .....
5601  GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT
      CGCGTCTTTT TTTCCTAGAG TTCTTCTAGG AAACTAGAAA AGATGCCCCA
 .....
5651  CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA
      GACTGCGAGT CACCTTGCTT TTGAGTGCAA TTCCCTAAAA CCAGTACTCT
 .....
5701  TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA AATGAAGTTT
      AATAGTTTTT CCTAGAAGTG GATCTAGGAA AATTTAATTT TTACTTCAAA
 .....
5751  TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT
      ATTTAGTTAG ATTTCATATA TACTCATTTG AACCAGACTG TCAATGGTTA
 .....
5801  GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC
      CGAATTAGTC ACTCCGTGGA TAGAGTCGCT AGACAGATAA AGCAAGTAGG
 .....
5851  ATAGTTGCCT GACTCCGCGG GGGGGGGGCG CTGAGGTCTG CCTCGTGAAG
      TATCAACGGA CTGAGGCCCC CCCCCCCCGC GACTCCAGAC GGAGCACTTC
 .....
5901  AAGGTGTTGC TGACTCATAC CAGGCCTGAA TCGCCCCATC ATCCAGCCAG
      TTCCACAACG ACTGAGTATG GTCCGGACTT AGCGGGGTAG TAGGTCGGTC
 .....
5951  AAAGTGAGGG AGCCACGGTT GATGAGAGCT TTGTTGTAGG TGGACCAGTT
      TTTCACTCCC TCGGTGCCAA CTACTCTCGA AACAACATCC ACCTGGTCAA
 .....
6001  GGTGATTTTG AACTTTTGCT TTGCCACGGA ACGGTCTGCG TTGTCGGGAA
      CCACTAAAAC TTGAAAACGA AACGGTGCCT TGCCAGACGC AACAGCCCTT
 .....
6051  GATGCGTGAT CTGATCCTTC AACTCAGCAA AAGTTCGATT TATTCAACAA
      CTACGCACTA GACTAGGAAG TTGAGTCGTT TTCAAGCTAA ATAAGTTGTT
 .....
6101  AGCCGCCGTC CCGTCAAGTC AGCGTAATGC TCTGCCAGTG TTACAACCAA
      TCGGCGGCAG CGCAGTTCAG TCGCATTACG AGACGGTCAC AATGTTGGTT
 .....
6151  TTAACCAATT CTGATTAGAA AAACTCATCG AGCATCAAAT GAAACTGCAA
      AATTGGTTAA GACTAATCTT TTTGAGTAGC TCGTAGTTTA CTTTGACGTT
 .....
6201  TTTATTCATA TCAGGATTAT CAATACCATA TTTTTGAAAA AGCCGTTTCT
      AAATAAGTAT AGTCCTAATA GTTATGGTAT AAAAACTTTT TCGGCAAAGA
 .....
```

```
6251  GTAATGAAGG AGAAAACTCA CCGAGGCAGT TCCATAGGAT GGCAAGATCC
      CATTACTTCC TCTTTTGAGT GGCTCCGTCA AGGTATCCTA CCGTTCTAGG

6301  TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC AACCTATTAA
      ACCATAGCCA GACGCTAAGG CTGAGCAGGT TGTAGTTATG TTGGATAATT

6351  TTTCCCCTCG TCAAAAATAA GGTTATCAAG TGAGAAATCA CCATGAGTGA
      AAAGGGGAGC AGTTTTTATT CCAATAGTTC ACTCTTTAGT GGTACTCACT

                                HindIII
                                ========
6401  CGACTGAATC CGGTGAGAAT GGCAAAAGCT TATGCATTTC TTTCCAGACT
      GCTGACTTAG GCCACTCTTA CCGTTTTCGA ATACGTAAAG AAAGGTCTGA

6451  TGTTCAACAG GCCAGCCATT ACGCTCGTCA TCAAAATCAC TCGCATCAAC
      ACAAGTTGTC CGGTCGGTAA TGCGAGCAGT AGTTTTAGTG AGCGTAGTTG

                                                      PvuI
                                                      ====
6501  CAAACCGTTA TTCATTCGTG ATTGCGCCTG AGCGAGACGA AATACGCGAT
      GTTTGGCAAT AAGTAAGCAC TAACGCGGAC TCGCTCTGCT TTATGCGCTA

      PvuI
      ==
6551  CGCTGTTAAA AGGACAATTA CAAACAGGAA TCGAATGCAA CCGGCGCAGG
      GCGACAATTT TCCTGTTAAT GTTTGTCCTT AGCTTACGTT GGCCGCGTCC

6601  AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG GATATTCTTC
      TTGTGACGGT CGCGTAGTTG TTATAAAAGT GGACTTAGTC CTATAAGAAG

6651  TAATACCTGG AATGCTGTTT TCCCGGGGAT CGCAGTGGTG AGTAACCATG
      ATTATGGACC TTACGACAAA AGGGCCCCTA GCGTCACCAC TCATTGGTAC

6701  CATCATCAGG AGTACGGATA AAATGCTTGA TGGTCGGAAG AGGCATAAAT
      GTAGTAGTCC TCATGCCTAT TTTACGAACT ACCAGCCTTC TCCGTATTTA

6751  TCCGTCAGCC AGTTTAGTCT GACCATCTCA TCTGTAACAT CATTGGCAAC
      AGGCAGTCGG TCAAATCAGA CTGGTAGAGT AGACATTGTA GTAACCGTTG

6801  GCTACCTTTG CCATGTTTCA GAAACAACTC TGGCGCATCG GGCTTCCCAT
      CGATGGAAAC GGTACAAAGT CTTTGTTGAG ACCGCGTAGC CCGAAGGGTA

6851  ACAATCGATA GATTGTCGCA CCTGATTGCC CGACATTATC GCGAGCCCAT
      TGTTAGCTAT CTAACAGCGT GGACTAACGG GCTGTAATAG CGCTCGGGTA

                                                      XhoI
                                                      ====
6901  TTATACCCAT ATAAATCAGC ATCCATGTTG GAATTTAATC GCGGCCTCGA
      AATATGGGTA TATTTAGTCG TAGGTACAAC CTTAAATTAG CGCCGGAGCT

      XhoI
      ====
6951  GCAAGACGTT TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT
      CGTTCTGCAA AGGGCAACTT ATACCGAGTA TTGTGGGGAA CATAATGACA

7001  TTATGTAAGC AGACAGTTTT ATTGTTCATG ATGATATATT TTTATCTTGT
      AATACATTCG TCTGTCAAAA TAACAAGTAC TACTATATAA AAATAGAACA
```

DraIII

```
7051  GCAATGTAAC ATCAGAGATT TTGAGACACA ACGTGGCTTT CCCCCCCCCC
      CGTTACATTG TAGTCTCTAA AACTCTGTGT TGCACCGAAA GGGGGGGGGG
 ..........................................................................
7101  CCATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC GGATACATAT
      GGTAATAACT TCGTAAATAG TCCCAATAAC AGAGTACTCG CCTATGTATA
 ..........................................................................
7151  TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCC
      AACTTACATA AATCTTTTTA TTTGTTTATC CCCAAGGCGC GTGTAAAGGG
 ..........................................................................
7201  CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAC
      GCTTTTCACG GTGGACTGCA GATTCTTTGG TAATAATAGT ACTGTAATTG
 ..........................................................................
7251  CTATAAAAAT AGGCGTATCA CGAGGCCCTT TCGTC
      GATATTTTTA TCCGCATAGT GCTCCGGGAA AGCAG
 ..........................................................................
```

pVR 1012-SGP(Z)

General Description

DNA pVR 1012-SGP(Z)
Local object
Created: 09/14/98 04:29PM
Last Modified: 09/15/98 04:50PM
length: 7272 bp
storage type: Basic
form: Circular
Comments

Restriction Map

DraIII: 1 site CACNNNGTG / GTGNNNCAC

HindIII: 1 site AAGCTT / TTCGAA

HpaI: 1 site GTTAAC / CAATTG

KpnI: 1 site GGTACC / CCATGG

NotI: 1 site GCGGCCGC / CGCCGGCG

PmlI: 1 site CACGTG / GTGCAC

PvuI: 1 site CGATCG / GCTAGC

SacII: 1 site CCGCGG / GGCGCC

XbaI: 1 site TCTAGA / AGATCT

XhoI: 1 site CTCGAG / GAGCTC

EcoRV: 2 sites GATATC / CTATAG

NcoI: 2 sites CCATGG / GGTACC

NdeI: 2 sites CATATG / GTATAC

SphI: 2 sites GCATGC / CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT
Start: 886 End: 1129

CMV IE INT
Start: 1130 End: 1840

TbGH
Start: 4289 End: 4841

Kanr
Start: 6337 End: 6959 (Complementary)

Misc_feature (2 signals)

CMV enhancer
Start: 248 End: 885
SGP(Z)
Start: 1870 End: 4288

Annotations

```
   1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

  51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
      CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

 101  TCAGGCGCGC TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
      AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                          NdeI
 151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
      GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACGC CACACTTTAT

 201  CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
      GGCGTGTCTA CGCATTCCTC TTTTATGGCG TAGTCTAACC GATAACCGGT

 251  TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
      AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

 301  TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
      AGGTTGTAAT GGCGGTACAA CTGTAACTAA TAACTGATCA ATAATTATCA

 351  AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
      TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA

 401  ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG
      TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC

 451  CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
      GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTGC GGTTATCCCT

 501  CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
      GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTG ACGGGTGAAC

                      NdeI
 551  GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
      CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT

 601  TGACGGTAAA TGGCCCGCCT GGCATTATGC CCAGTACATG ACCTTATGGG
      ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGGAATACCC

                                                        NcoI
 651  ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
      TGAAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

      NcoI
 701  GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
      CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG

 751  ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
      TGCCCCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAAA

 801  GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
      CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT
```

```
 851  TTCACGCAAA TGGGCGGTAG GCGTGTACGG TGGGAGGTCT ATATAAGCAG
      AAGTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC
 ..........................................................................

 901  AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
      TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA
 ..........................................................................

                                                  SacII
                                                 ========

 951  TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
      AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT
 ..........................................................................

1001  CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
      GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG
 ..........................................................................

                                                  SphI
                                                 ========

1051  CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
      GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA
 ..........................................................................

1101  TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
      AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC
 ..........................................................................

1151  TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
      ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG
 ..........................................................................

1201  TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
      ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT
 ..........................................................................

1251  CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
      GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG
 ..........................................................................

1301  ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
      TGCCTGAGAC ATAAAAATGT CCTACCCCAG GGTAAATAAT AAATGTTTAA
 ..........................................................................

1351  CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
      GTGTATATGT TGTTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT
 ..........................................................................

1401  GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
      CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA
 ..........................................................................

1451  TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
      AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG
 ..........................................................................

1501  AGCGGCTCAT GGTCGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
      TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC
 ..........................................................................

1551  ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
      TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC
 ..........................................................................

1601  TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
      ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGAGCGTGC
 ..........................................................................

1651  GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
      CGACTGCGTC TACCTTCTGA ATTCCGTCGC CGTCTTCTTC TACGTCCGTC
 ..........................................................................

1701  CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
      GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG
 ..........................................................................
```

```
        HpaI
1751  TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
      ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

                                                        NcoI
1801  CGCCCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
      GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA

      NcoI                              PmlI          EcoRVNotI
1851  GGGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
      CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

      NotI XbaI
1901  CCGCTCTAGA CCAGGCGCCT GGATCGAATT GATGAAGATT AAGCCGACAG
      GGCGAGATCT GGTCCGCGGA CCTAGCTTAA CTACTTCTAA TTCGGCTGTC

1951  TGAGCGTAAT CTTCATCTCT CTTAGATTAT TTGTTTTCCA GAGTAGGGGT
      ACTCGCATTA GAAGTAGAGA GAATCTAATA AACAAAAGGT CTCATCCCCA

2001  CGTCAGGTCC TTTTCAATCG TGTAACCAAA ATAAACTCCA CTAGAAGGAT
      GCAGTCCAGG AAAAGTTAGC ACATTGGTTT TATTTGAGGT GATCTTCCTA

2051  ATTGTGGGGC AACAACACAA TGGGCGTTAC AGGAATATTG CAGTTACCTC
      TAACACCCCG TTGTTGTGTT ACCCGCAATG TCCTTATAAC GTCAATGGAG

2101  GTGATCGATT CAAGAGGACA TCATTCTTTC TTTGGGTAAT TATCCTTTTC
      CACTAGCTAA GTTCTCCTGT AGTAAGAAAG AAACCCATTA ATAGGAAAAG

2151  CAAAGAACAT TTTCCATCCC ACTTGGAGTC ATCCACAATA GCACATTACA
      GTTTCTTGTA AAAGGTAGGG TGAACCTCAG TAGGTGTTAT CGTGTAATGT

2201  GGTTAGTGAT GTCGACAAAC TAGTTTGTCG TGACAAACTG TCATCCACAA
      CCAATCACTA CAGCTGTTTG ATCAAACAGC ACTGTTTGAC AGTAGGTGTT

2251  ATCAATTGAG ATCAGTTGGA CTGAATCTCG AAGGGAATGG AGTGGCAACT
      TAGTTAACTC TAGTCAACCT GACTTAGAGC TTCCCTTACC TCACCGTTGA

2301  GACGTCCCAT CTGCAACTAA AAGATGGGGC TTCAGGTCCG GTGTCCCACC
      CTGCAGGGTA GACGTTGATT TTCTACCCCG AAGTCCAGGC CACAGGGTGG

2351  AAAGGTGGTC AATTATGAAG CTGGTGAATG GGCTGAAAAC TGCTACAATC
      TTTCCACCAG TTAATACTTC GACCACTTAC CCGACTTTTG ACGATGTTAG

2401  TTGAAATCAA AAAACCTGAC GGGAGTGAGT GTCTACCAGC AGCGCCAGAC
      AACTTTAGTT TTTTGGACTG CCCTCACTCA CAGATGGTCG TCGCGGTCTG

2451  CCGATTCGGG GCTTCCCCCG GTGCCGGTAT GTGCACAAAG TATCAGGAAC
      GGCTAAGCCC CGAAGGGGGC CACGGCCATA CACGTGTTTC ATAGTCCTTG

2501  GGCACCGTGT GCCGGAGACT TTGCCTTCCA TAAAGAGGGT GCTTTCTTCC
      CCCTGGCACA CGGCCTCTGA AACGGAAGGT ATTTCTCCCA CGAAAGAAGG

2551  TGTATGATCG ACTTGCTTCC ACAGTTATCT ACCGAGGAAC GACTTTCGCT
      ACATACTAGC TGAACGAAGG TGTCAATAGA TGGCTCCTTG CTGAAAGCGA
```

```
2601  GAAGGTCTCG TTGCATTTCT GATACTGCCC CAAGCTAAGA AGGACTTCTT
      CTTCCAGAGC AACGTAAAGA CTATGACGGG GTTCGATTCT TCCTGAAGAA
 .....

2651  CAGCTCACAC CCCTTGAGAG AGCCGGTCAA TGCAACGGAG GACCCCTCTA
      GTCGAGTGTG GGGAACTCTC TCGGCCAGTT ACGTTGCCTC CTGGGCAGAT
 .....

                                  EcoRV
                                  ------

2701  GTGGCTACTA TTCTACCACA ATTAGATATC AGGCTACCGG TTTTGGAACC
      CACCGATGAT AAGATGGTGT TAATCTATAG TCCGATGGCC AAAACCTTGG
 .....

2751  AATGACACAG AGTACTTGTT CGAGGTTGAC AATTTGACCT ACGTCCAACT
      TTACTCTGTC TCATGAACAA GCTCCAACTG TTAAACTGGA TGCAGGTTGA
 .....

2801  TGAATCAAGA TTCACACCAC AGTTTCTGCT CCAGCTGAAT GAGACAATAT
      ACTTCAGTTCT AAGTGTGGTG TCAAAGACGA GGTCGACTTA CTCTGTTATA
 .....

2851  ATACAAGTGG GAAAAGGAGC AATACCACGG GAAAACTAAT TTGGAAGGTC
      TATGTTCACC CTTTTCCTCG TTATGGTGCC CTTTTGATTA AACCTTCCAG
 .....

2901  AACCCCGAAA TTGATACAAC AATCGGGGAG TGGGCCTTCT GGGAAACTAA
      TTGGGGCTTT AACTATGTTG TTAGCCCCTC ACCCGGAAGA CCCTTTGATT
 .....

2951  AAAAACCTCA CTAGAAAAAT TCGCAGTGAA GAGTTGTCTT TCACAGTTGT
      TTTTTGGAGT GATCTTTTTA AGCGTCACTT CTCAACAGAA AGTGTCAACA
 .....

3001  ATCAAACGGA GCCAAAAACA TCAGTGGTCA GAGTCCGGCG CGAACTTCTT
      TAGTTTGCCT CGGTTTTTGT AGTCACCAGT CTCAGGCCGC GCTTGAAGAA
 .....

3051  CCGACCCAGG GACCAACACA ACAACTGAAG ACCACAAAAT CATGGCTTCA
      GGCTGGGTCC CTGGTTGTGT TGTTGACTTC TGGTGTTTTA GTACCGAAGT
 .....

3101  GAAAATTCCT CTGCAATGGT TCAAGTGCAC AGTCAAGGAA GGGAAGCTGC
      CTTTTAAGGA GACGTTACCA AGTTCACGTG TCAGTTCCTT CCCTTCGACG
 .....

3151  AGTGTCGCAT CTAACAACCC TTGCCACAAT CTCCACGAGT CCCCAATCCC
      TCACAGCGTA GATTGTTGGG AACGGTGTTA GAGGTGCTCA GGGGTTAGGG
 .....

3201  TCACAACCAA ACCAGGTCCG GACAACAGCA CCCATAATAC ACCCGTGTAT
      AGTGTTGGTT TGGTCCAGGC CTGTTGTCGT GGGTATTATG TGGGCACATA
 .....

3251  AAACTTGACA TCTCTGAGGC AACTCAAGTT GAACAACATC ACCGCAGAAC
      TTTGAACTGT AGAGACTCCG TTGAGTTCAA CTTGTTGTAG TGGCGTCTTG
 .....

3301  AGACAACGAC AGCACAGCCT CCGACACTCC CTCTGCCACG ACCGCAGCCG
      TCTGTTGCTG TCGTGTCGGA GGCTGTGAGG GAGACGGTGC TGGCGTCGGC
 .....

3351  GACCCCCAAA AGCAGAGAAC ACCAACACGA GCAAGAGCAC TGACTTCCTC
      CTGGGGGTTT TCGTCTCTTG TGGTTGTGCT CGTTCTCGTG ACTGAAGGAC
 .....

3401  GACCCCGCCA CCACAACAAG TCCCCAAAAC CACAGCGAGA CCGCTGGCAA
      CTGGGGCGGT GGTGTTGTTC AGGCGTTTTG GTGTCGCTCT GGCGACCGTT
 .....

3451  CAACAACACT CATCACCAAG ATACCGGAGA AGAGAGTGCC AGCAGCGGGA
      GTTGTTGTGA GTAGTGGTTC TATGGCCTCT TCTCTCACGG TCGTCGCCCT
 .....
```

```
3501 AGCTAGGCTT AATTAGCAAT ACTATTGCTG GAGTCGCAGG ACTGATCACA
     TCGATCCGAA TTAATCGTTA TGATAACGAC CTCAGCGTCC TGACTAGTGT

3551 GGCGGGAGAA GAACTCGAAG AGAAGCAATT GTCAATGCTC AACCCAAATG
     CCGCCCTCTT CTTGAGCTTC TCTTCGTTAA CAGTTACGAG TTGGGTTTAC

3601 CAACCCTAAT TTACATTACT GGACTACTCA GGATGAAGGT GCTGCAATCG
     GTTGGGATTA AATGTAATGA CCTGATGAGT CCTACTTCCA CGACGTTAGC

3651 GACTGGCCTG GATACCATAT TTCGGGCCAG CAGCCGAGGG AATTTACATA
     CTGACCGGAC CTATGGTATA AAGCCCGGTC GTCGGCTCCC TTAAATGTAT

3701 GAGGGGCTAA TGCACAATCA AGATGGTTTA ATCTGTGGGT TGAGACAGCT
     CTCCCCGATT ACGTGTTAGT TCTACCAAAT TAGACACCCA ACTCTGTCGA

3751 GGCCAACGAG ACGACTCAAG CTCTTCAACT GTTCCTGAGA GCCACAACTG
     CCGGTTGCTC TGCTGAGTTC GAGAAGTTGA CAAGGACTCT CGGTGTTGAC

3801 AGCTACGCAC CTTTTCAATC CTCAACCGTA AGGCAATTGA TTTCTTGCTG
     TCGATGCGTG GAAAAGTTAG GAGTTGGCAT TCCGTTAACT AAAGAACGAC

3851 CAGCGATGGG GCGGCACATG CCACATTCTG GGACCGGACT GCTGTATCGA
     GTCGCTACCC CGCCGTGTAC GGTGTAAGAC CCTGGCCTGA CGACATAGCT

3901 ACCACATGAT TGGACCAAGA ACATAACAGA CAAAATTGAT CAGATTATTC
     TGGTGTACTA ACCTGGTTCT TGTATTGTCT GTTTTAACTA GTCTAATAAG

3951 ATGATTTTGT TGATAAAACC CTTCCGGACC AGGGGGACAA TGACAATTGG
     TACTAAAACA ACTATTTTGG GAAGGCCTGG TCCCCCTGTT ACTGTTAACC

4001 TGGACAGGAT GGAGACAATG GATACCGGCA GGTATTGGAG TTACAGGCGT
     ACCTGTCCTA CCTCTGTTAC CTATGGCCGT CCATAACCTC AATGTCCGCA

4051 TATAATTGCA GTTATCGCTT TATTCTGTAT ATGCAAATTT GTCTTTTAGT
     ATATTAACGT CAATAGCGAA ATAAGACATA TACGTTTAAA CAGAAAATCA

4101 TTTTCTTCAG ATTGCTTCAT GGAAAAGCTC AGCCTCAAAT CAATGAAACC
     AAAAGAAGTC TAACGAAGTA CCTTTTCGAG TCGGAGTTTA GTTACTTTGG

4151 AGGATTTAAT TATATGGATT ACTTGAATCT AAGATTACTT GACAAATGAT
     TCCTAAATTA ATATACCTAA TGAACTTAGA TTCTAATGAA CTGTTTACTA

4201 AATATAATAC ACTGGAGCTT TAAACATAGC CAATGTGATT CTAACTCCTT
     TTATATTATG TGACCTCGAA ATTTGTATCG GTTACACTAA GATTGAGGAA

4251 TAAACTCACA GTTAATCATA AACAAGGTTT GGAATTGATC TGCTGTGCCT
     ATTTGAGTGT CAATTAGTAT TTGTTCCAAA CCTTAACTAG ACGACACGGA

4301 TCTAGTTGCC AGCCATCTGT TGTTTGCCCC TCCCCCGTGC CTTCCTTGAC
     AGATCAACGG TCGGTAGACA ACAAACGGGG AGGGGGCACG GAAGGAACTG

4351 CCTGGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
     GGACCTTCCA CGGTGAGGGT GACAGGAAAG GATTATTTTA CTCCTTTAAC

4401 CATCGCATTG TCTGAGTAGG TGTCATTCTA TTCTGGGGGG TGGGGTGGGG
     GTAGCGTAAC AGACTCATCC ACAGTAAGAT AAGACCCCCC ACCCCACCCC
```

SphI

4451 CACCACAGCA AGGGCGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA
GTCGTGTCGT TCCCCCTCCT AACCCTTCTG TTATCGTCCG TACGACCCCT

KpnI

4501 TGCGGTGGGC TCTATGGGTA CCCAGGTGCT GAAGAATTGA CCCGGTTCCT
ACGCCACCCG AGATACCCAT GGGTCCACGA CTTCTTAACT GGGCCAAGGA

4551 CCTGGGCCAG AAAGAAGCAG GCACATCCCC TTCTCTGTGA CACACCCTGT
GGACCCGGTC TTTCTTCGTC CGTGTAGGGG AAGAGACACT GTGTGGGACA

4601 CCACCCCCCT CGTTCTTAGT TCCAGCCCCA CTCATAGGAC ACTCATAGCT
GGTGCCGGGA CCAAGAATCA AGGTCGGGGT GAGTATCCTG TGAGTATCGA

4651 CAGGAGGCCT CCGCCTTCAA TCCCACCCGC TAAAGTACTT GGAGCGGTCT
GTCCTCCCGA GGCGGAAGTT AGGGTGGGCG ATTTCATGAA CCTCGCCAGA

4701 CTCCCTCCCT CATCAGCCCA CCAAACCAAA CCTAGCCTCC AAGAGTGGGA
GAGGGAGGGA GTAGTCGGGT GGTTTGGTTT GGATCGGAGG TTCTCACCCT

4751 AGAAACTAAA GCAAGATAGG CTATTAAGTC CAGAGGCAGA GAAAATGCCT
TCTTTAATTT CGTTCTATCC GATAATTCAG GTCTCCCTCT CTTTTACGGA

4801 CCAACATGTG AGGAACTAAT GAGAGAAATC ATAGAATTTC TTCCGCTTCC
GGTTGTACAC TCCTTCATTA CTCTCTTTAG TATCTTAAAG AAGGCGAAGG

4851 TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC GAGCGGTATC
AGCGAGTGAC TGAGCGACGC GAGCCAGCAA GCCGACGCCG CTCGCCATAG

4901 AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG
TCGAGTGAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT CCCCTATTGC

4951 CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA
GTCCTTTCTT GTACACTCGT TTTCCGGTCG TTTTCCGGTC CTTGGCATTT

5001 AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC CTGACGAGCA
TTCCGGCGCA ACGACCCCAA AAAGGTATCC GAGGCGGGGG GACTGCTCGT

5051 TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
AGTGTTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC TGTCCTGATA

5101 AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT
TTTCTATGGT CCCCAAAGGG GGACCTTCGA GGGAGCACGC GAGAGGACAA

5151 CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCGGGAAG
GGCTGCGACG CCGAATGGCC TATGGACAGG CGGAAAGAGG GAAGCCCTTC

5201 CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
GCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA AGCCACATCC

5251 TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC
AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA AGTCGGGCTG

5301 CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA
GCGACGCGGA ATAGGCCATT GATAGCAGAA CTCAGGTTGG GCCATTCTGT

```
5351  CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA
      GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCCTAA TCGTCTCGCT
 .....

5401  GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACTACGGC
      CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG ATTGATGCCG
 .....

5451  TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC
      ATGTGATCTT CCTGTCATAA ACCATAGACG CGAGACGACT TCGGTCAATG
 .....

5501  CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG
      GAAGCCTTTT TCTCAACCAT CGAGAACTAG GCCGTTTGTT TGGTGGCGAC
 .....

5551  GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA
      CATCGCCACC AAAAAAACAA ACGTTCGTCG TCTAATGCGC GTCTTTTTTT
 .....

5601  GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG ACGCTCAGTG
      CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC TGCGAGTCAC
 .....

5651  GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA TCAAAAAGGA
      CTTGCTTTTG AGTGCAATTC CCTAAAACCA GTACTCTAAT AGTTTTTCCT
 .....

5701  TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA ATCAATCTAA
      AGAAGTGGAT CTAGGAAAAT CTAATTTTTA CTTCAAAATT TAGTTAGATT
 .....

5751  AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA
      TCATATATAC TCATTTGAAC CAGACTGTCA ATGGTTACGA ATTAGTCACT
 .....

5801  GGCACCTATC TCAGCGATCT GTCTATTTCG TTCATCCATA GTTGCCTGAC
      CCGTGGATAG AGTCGCTAGA CAGATAAAGC AAGTAGGTAT CAACGGACTG
 .....

5851  TCCCCGGGGG GGGGCCGCTG AGGTCTGCCT CGTGAAGAAG GTGTTGCTGA
      AGGCCCCCCC CCCCCGCGAC TCCAGACGGA GCACTTCTTC CACAACGACT
 .....

5901  CTCATACCAG GCCTGAATCG CCCCATCATC CAGCCAGAAA GTGAGGGAGC
      GAGTATGGTC CGGACTTAGC GGGGTAGTAG GTCGGTCTTT CACTCCCTCG
 .....

5951  CACGGTTGAT GAGAGCTTTG TTGTAGGTGG ACCAGTTGGT GATTTTGAAC
      GTGCCAACTA CTCTCGAAAC AACATCCACC TGGTCAACCA CTAAAACTTG
 .....

6001  TTTTGCTTTG CCACGGAACG GTCTGCGTTG TCGGGAAGAT GCGTGATCTG
      AAAACGAAAC GGTGCCTTGC CAGACGCAAC AGCCCTTCTA CGCACTAGAC
 .....

6051  ATCCTTCAAC TCAGCAAAAG TTCGATTTAT TCAACAAAGC CGCCGTCCCG
      TAGGAAGTTG AGTCGTTTTC AAGCTAAATA AGTTGTTTCG GCGGCAGGGC
 .....

6101  TCAAGTCAGC GTAATGCTCT GCCAGTGTTA CAACCAATTA ACCAATTCTG
      AGTTCAGTCG CATTACGAGA CGGTCACAAT GTTGGTTAAT TGGTTAAGAC
 .....

6151  ATTAGAAAAA CTCATCGAGC ATCAAATGAA ACTGCAATTT ATTCATATCA
      TAATCTTTTT GAGTAGCTCG TAGTTTACTT TGACGTTAAA TAAGTATAGT
 .....

6201  GCATTATCAA TACCATATTT TTGAAAAAGC CGTTTCTGTA ATGAAGGAGA
      CGTAATAGTT ATGGTATAAA AACTTTTTCG GCAAAGACAT TACTTCCTCT
 .....

6251  AAACTCACCG AGGCAGTTCC ATAGGATGGC AAGATCCTGG TATCGGTCTG
      TTTGAGTGGC TCCGTCAAGG TATCCTACCG TTCTAGGACC ATAGCCAGAC
 .....
```

```
6301  CGATTCCGAC TCGTCCAACA TCAATACAAC CTATTAATTT CCCCTCGTCA
      GCTAAGGCTG AGCAGGTTGT AGTTATGTTG GATAATTAAA GGGGAGCAGT
.....
6351  AAAATAAGGT TATCAAGTGA GAAATCACCA TGAGTGACGA CTGAATCCGG
      TTTTATTCCA ATAGTTCACT CTTTAGTGGT ACTCACTGCT GACTTAGGCC
.....
                HindIII
                ========
6401  TGAGAATGGC AAAAGCTTAT GCATTTCTTT CCAGACTTGT TCAACAGGCC
      ACTCTTACCG TTTTCGAATA CGTAAAGAAA GGTCTGAACA AGTTGTCCGG
.....
6451  AGCCATTACG CTCGTCATCA AAATCACTCG CATCAACCAA ACCGTTATTC
      TCGGTAATGC GAGCAGTAGT TTTAGTGAGC GTAGTTGGTT TGGCAATAAG
.....
                                        PvuI
                                        ====
6501  ATTCGTGATT GCGCCTGAGC GAGACGAAAT ACGCGATCGC TGTTAAAAGG
      TAAGCACTAA CGCGGACTCG CTCTGCTTTA TGCGCTAGCG ACAATTTTCC
.....
6551  ACAATTACAA ACAGGAATCG AATGCAACCG GCGCAGGAAC ACTGCCAGCG
      TGTTAATGTT TGTCCTTAGC TTACGTTGGC CGCGTCCTTG TGACGGTCGC
.....
6601  CATCAACAAT ATTTTCACCT GAATCAGGAT ATTCTTCTAA TACCTGGAAT
      GTAGTTGTTA TAAAAGTGGA CTTAGTCCTA TAAGAAGATT ATGGACCTTA
.....
6651  GCTGTTTTCC CGGGGATCGC AGTGGTGAGT AACCATGCAT CATCAGGAGT
      CGACAAAAGG GCCCCTAGCG TCACCACTCA TTGGTACGTA GTAGTCCTCA
.....
6701  ACGGATAAAA TGCTTGATGG TCGGAAGAGG CATAAATTCC GTCAGCCAGT
      TGCCTATTTT ACGAACTACC AGCCTTCTCC GTATTTAAGG CAGTCGGTCA
.....
6751  TTAGTCTGAC CATCTCATCT GTAACATCAT TGGCAACGCT ACCTTTGCCA
      AATCAGACTG GTAGAGTAGA CATTGTAGTA ACCGTTGCGA TGGAAACGGT
.....
6801  TGTTTCAGAA ACAACTCTGG CGCATCGGGC TTCCCATACA ATCGATAGAT
      ACAAAGTCTT TGTTGAGACC GCGTAGCCCG AAGGGTATGT TAGCTATCTA
.....
6851  TGTCGCACCT GATTGCCCGA CATTATCGCG AGCCCATTTA TACCCATATA
      ACAGCGTGGA CTAACGGGCT GTAATAGCGC TCGGGTAAAT ATGGGTATAT
.....
                                        XhoI
                                        ====
6901  AATCAGCATC CATGTTGGAA TTTAATCGCG GCCTCGAGCA AGACGTTTCC
      TTAGTCGTAG GTACAACCTT AAATTAGCGC CGGAGCTCGT TCTGCAAAGG
.....
6951  CGTTGAATAT GGCTCATAAC ACCCCTTGTA TTACTGTTTA TGTAAGCAGA
      GCAACTTATA CCGAGTATTG TGGGGAACAT AATGACAAAT ACATTCGTCT
.....
7001  CAGTTTTATT GTTCATGATG ATATATTTTT ATCTTGTGCA ATGTAACATC
      GTCAAAATAA CAAGTACTAC TATATAAAAA TAGAACACGT TACATTGTAG
.....
                DraIII
                ======
7051  AGAGATTTTG AGACACAACG TGGCTTTCCC CCCCCCCCCA TTATTGAAGC
      TCTCTAAAAC TCTGTGTTGC ACCGAAAGGG GGGGGGGGGT AATAACTTCG
.....
7101  ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA
      TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT
.....
```

```
7151  CAAAAACAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC
      GTTTTTATTT GTTTATCCCC AAGGCGCGTG TAAAGGGGCT TTTCACGGTG
.....
7201  CTGACGTCTA AGAAACCATT ATTATCATGA CATTAACCTA TAAAAATAGG
      GACTGCAGAT TCTTTGGTAA TAATAGTACT GTAATTGGAT ATTTTTATCC
.....
7251  CGTATCACGA GGCCCTTTCG TC
      GCATAGTGCT CCGGGAAAGC AG
.....
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/01382

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 48/00; C07H 21/04; C12N 15/63, 15/86, 5/10, 15/40
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/450, 93.2, 93.21; 536/23.1, 23.72; 435/5, 6, 455, 457, 458, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog, Biosis, Medline, Biotech
Search terms: Ebola virus, glycoprotein, transmembrane, targeting

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,166,320 A (WU et al.) 24 November 1992, columns 9-10 and claims 1-18.	1-20
A	SANCHEZ et al. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. Proceedings of the National Academy of Sciences. April 1996, Vol. 93, pages 3602-3607, especially page 3604.	3, 12
A	FELGNER et al. Lipofection: A highly efficient, lipid-mediated DNA-transfection procedure. Proceedings of the National Academy of Sciences. November 1987, Vol. 84, pages 7413-7417, especially page 7414.	1-20

[X] Further documents are listed in the continuation of Box C [ ] See patent family annex.

*	Special categories of cited documents	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 12 APRIL 1999	Date of mailing of the international search report 11 MAY 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer DAVID GUZO Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/01382

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	VOLCHKOV et al. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. Virology. 1995, Vol. 214, pages 421-430, especially page 424.	3, 12

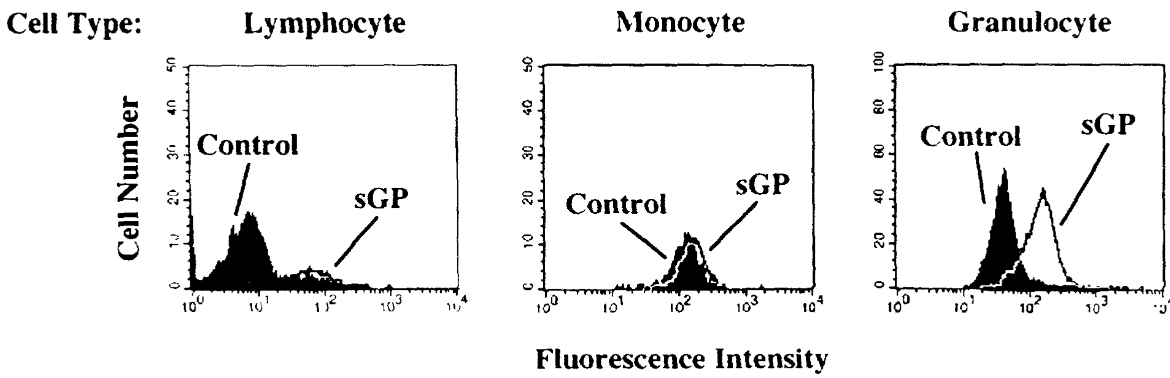
INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/01382

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

424/450, 93.2, 93.21; 536/23.1, 23.72; 435/5, 6, 455, 457, 458, 320.1

Figure 1A **Figure 1A1** **Figure 1A2**



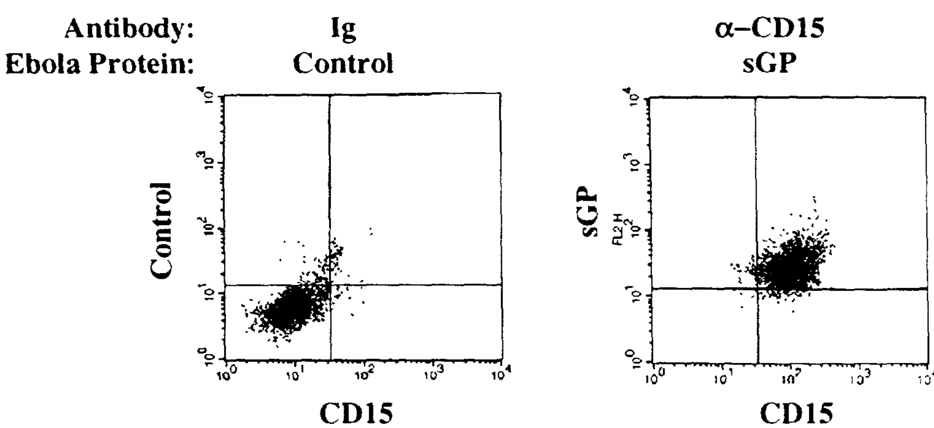


Figure 1B **Figure 1B1**

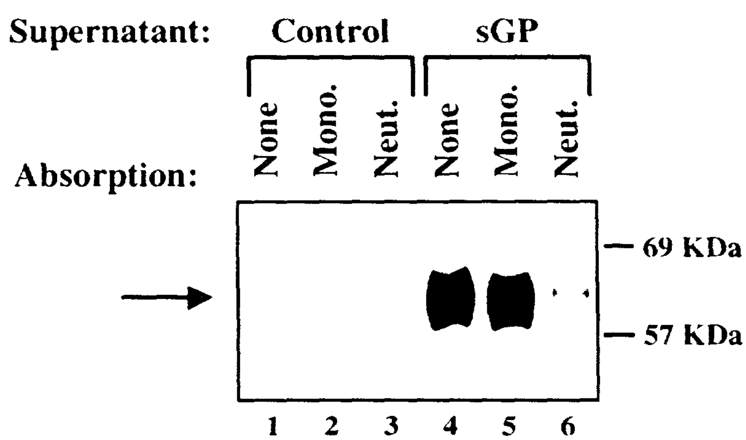


Figure 1C

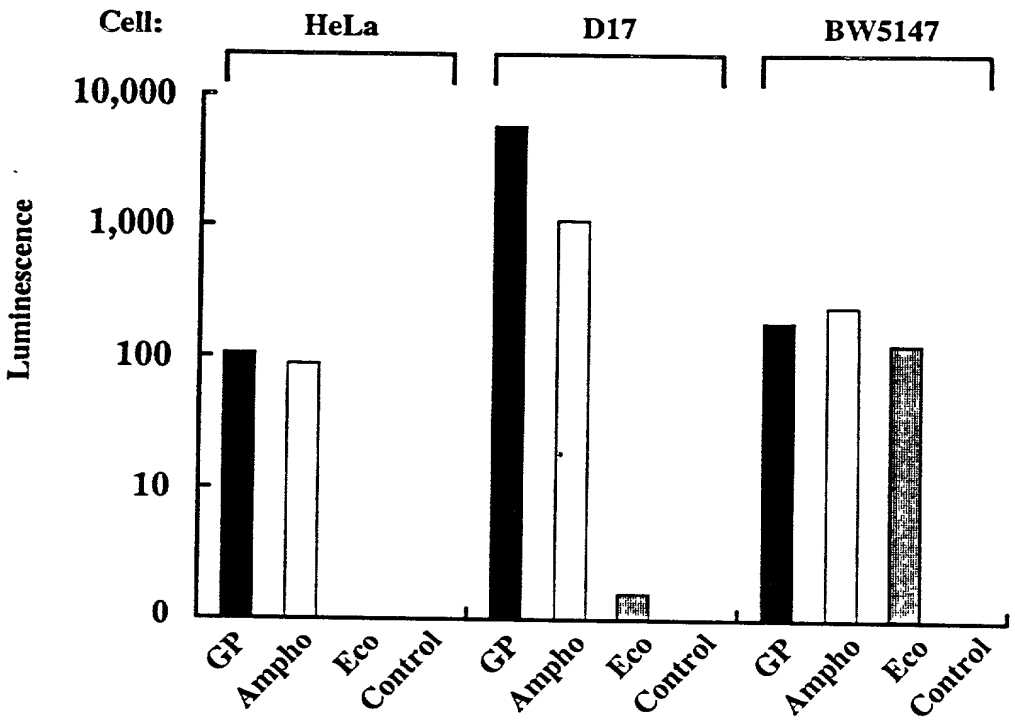


Figure 2A

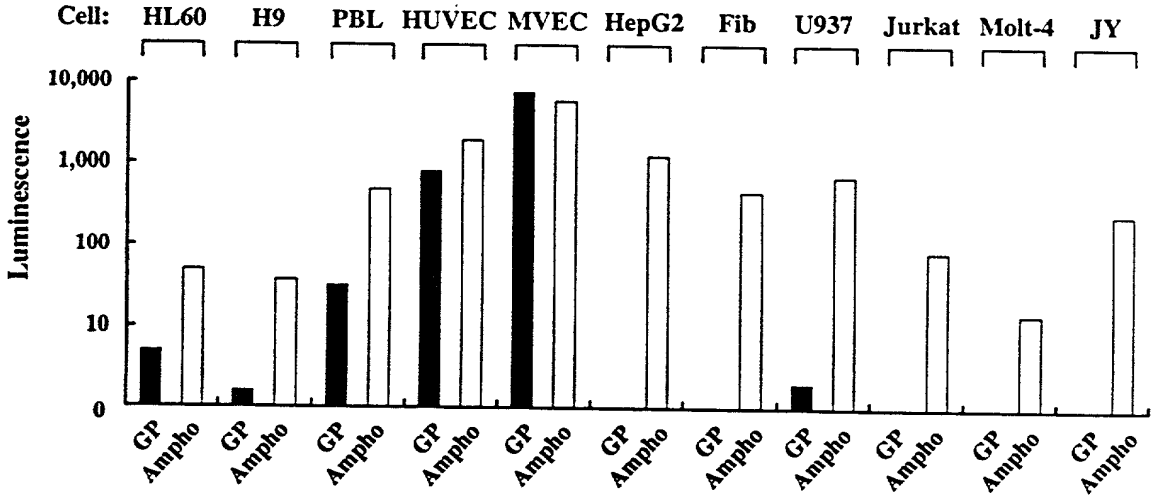


Figure 2B

Figure 2C

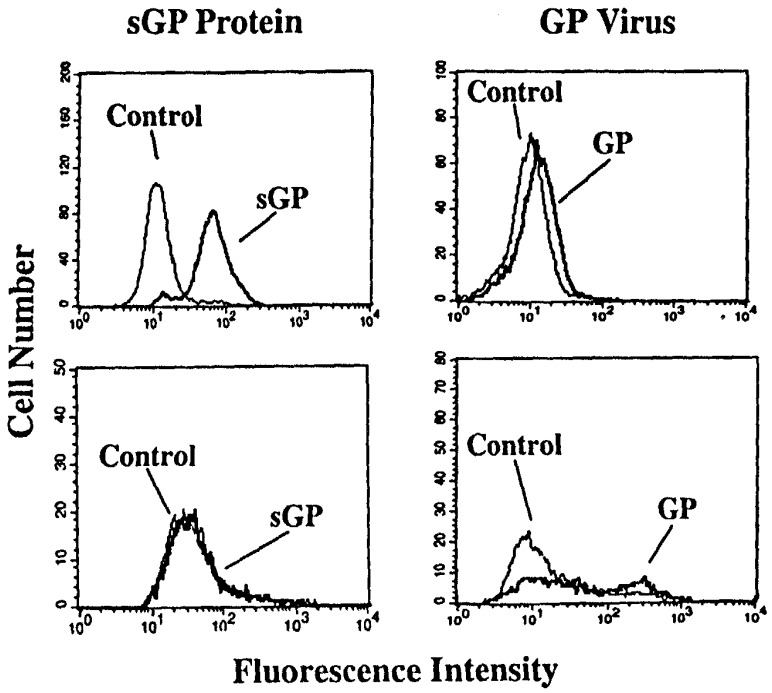


Figure 2C1

Figure 2C2

Figure 2C3

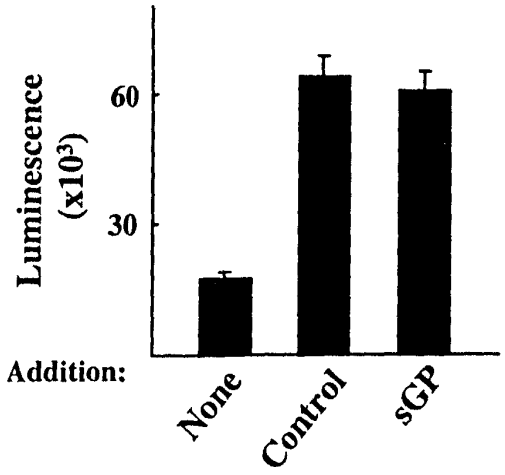


Figure 2D

Figure 3A

Ab: Control (Ig)

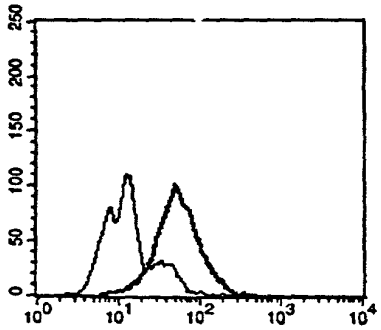


Figure 3B

α-CD16

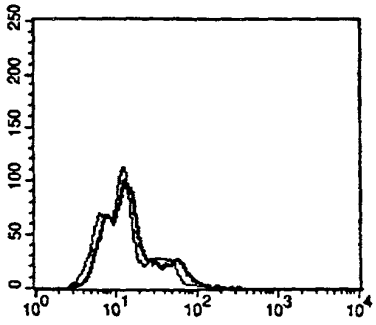


Figure 3C

Ab: Control (Ig)

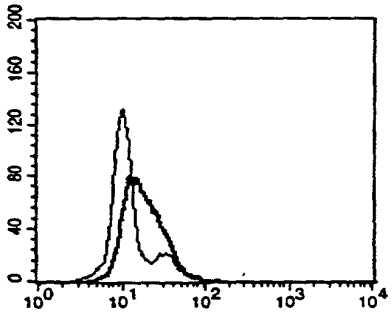


Figure 3D

α-CD62L

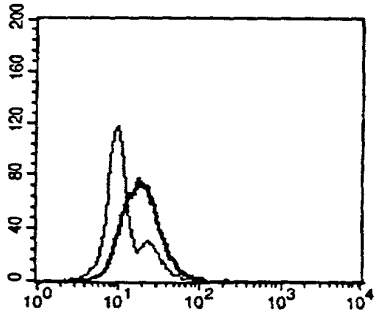


Figure 3E

Stimulation: None

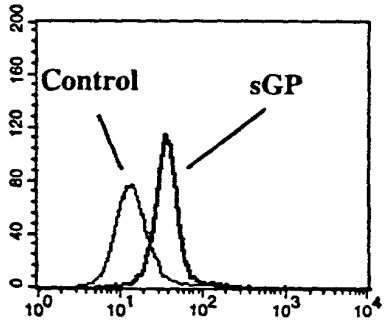
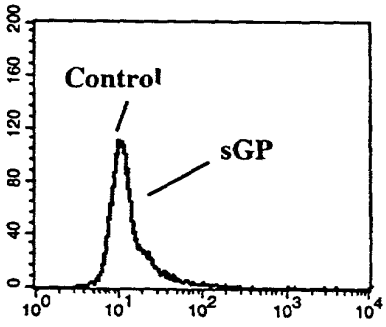


Figure 3F

PMA



Fluorescence Intensity

Figure 4A

Pre-stimulation

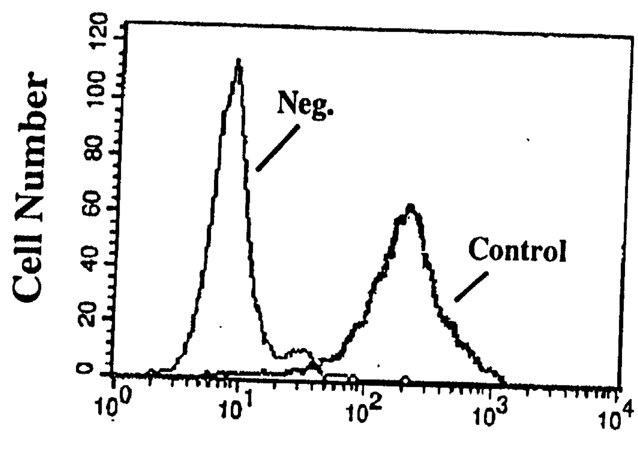
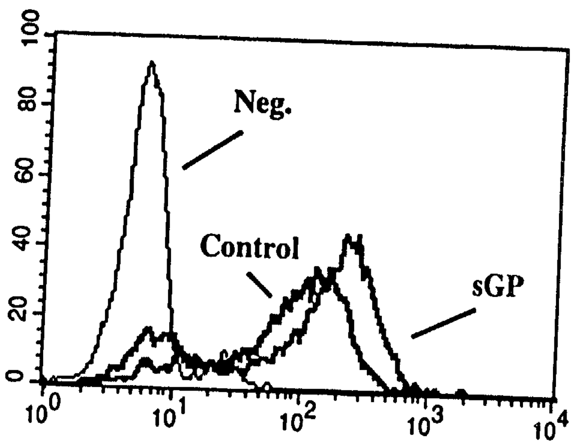


Figure 4B

Post-stimulation



Fluorescence Intensity

Figure 5A

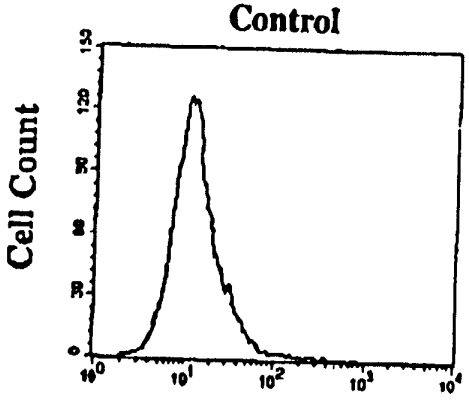


Figure 5B

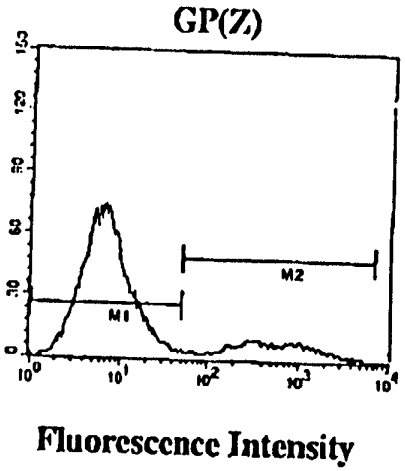
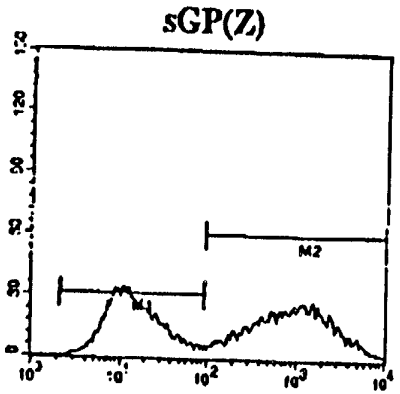


Figure 5C



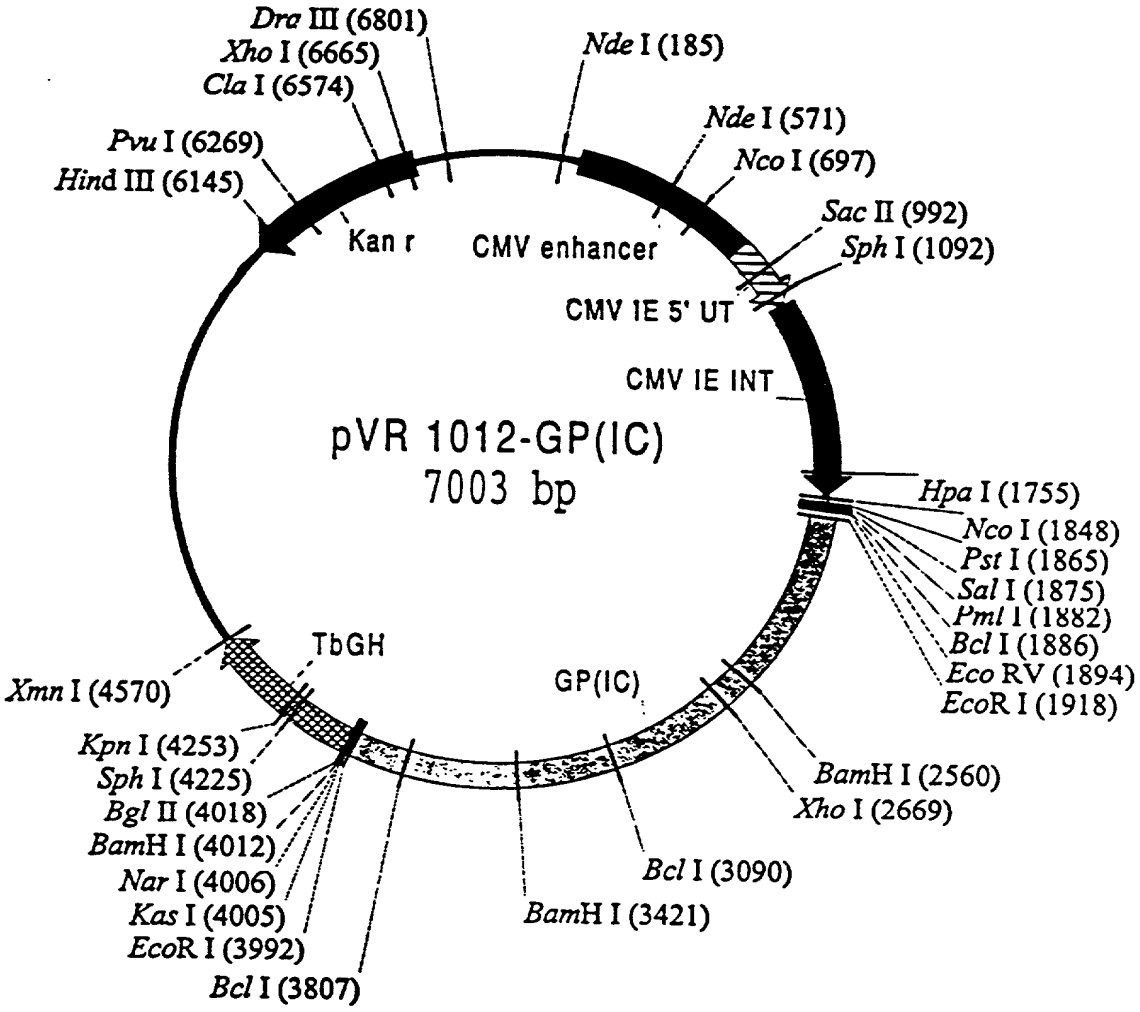

Dra III (6801)
Xho I (6665)
Cla I (6574)
Pvu I (6269)
Hind III (6145)
Nde I (185)
Nde I (571)
Nco I (697)
Sac II (992)
Sph I (1092)
Kan r
CMV enhancer
CMV IE 5' UT
CMV IE INT
pVR 1012-GP(IC)
7003 bp
Hpa I (1755)
Nco I (1848)
Pst I (1865)
Sal I (1875)
Pml I (1882)
Bcl I (1886)
Eco RV (1894)
EcoR I (1918)
BamH I (2560)
Xho I (2669)
Bcl I (3090)
BamH I (3421)
GP(IC)
TbGH
Xmn I (4570)
Kpn I (4253)
Sph I (4225)
Bgl II (4018)
BamH I (4012)
Nar I (4006)
Kas I (4005)
EcoR I (3992)
Bcl I (3807)

Figure 6

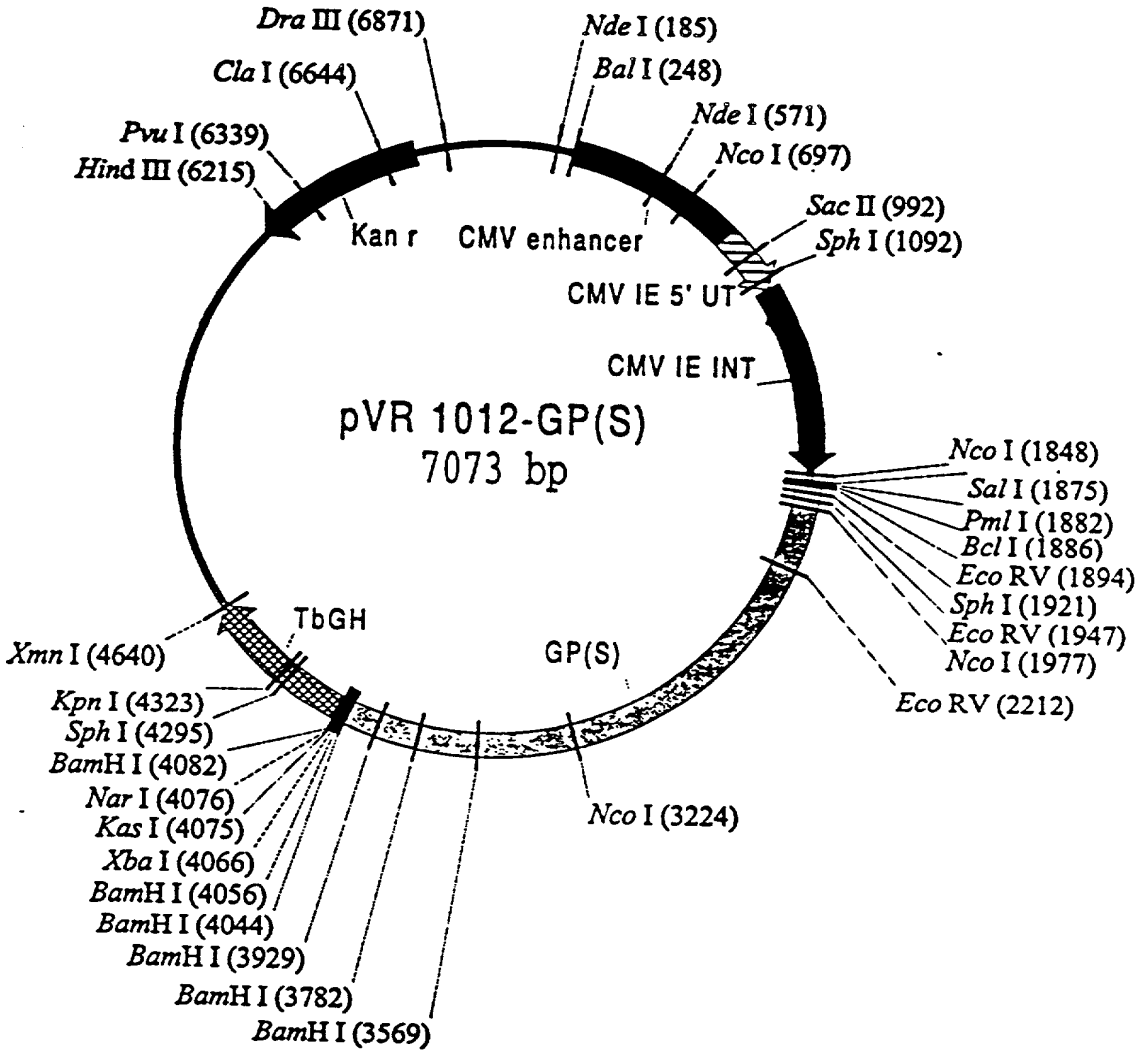

Dra III (6871)
Nde I (185)
Cla I (6644)
Bal I (248)
Nde I (571)
Pvu I (6339)
Nco I (697)
Hind III (6215)
Sac II (992)
Kan r
CMV enhancer
Sph I (1092)
CMV IE 5' UT
CMV IE INT
pVR 1012-GP(S)
7073 bp
Nco I (1848)
Sal I (1875)
Pml I (1882)
Bcl I (1886)
Eco RV (1894)
Sph I (1921)
Eco RV (1947)
Nco I (1977)
Eco RV (2212)
TbGH
Xmn I (4640)
GP(S)
Kpn I (4323)
Sph I (4295)
BamH I (4082)
Nar I (4076)
Kas I (4075)
Xba I (4066)
BamH I (4056)
BamH I (4044)
BamH I (3929)
BamH I (3782)
BamH I (3569)
Nco I (3224)

Figure 7

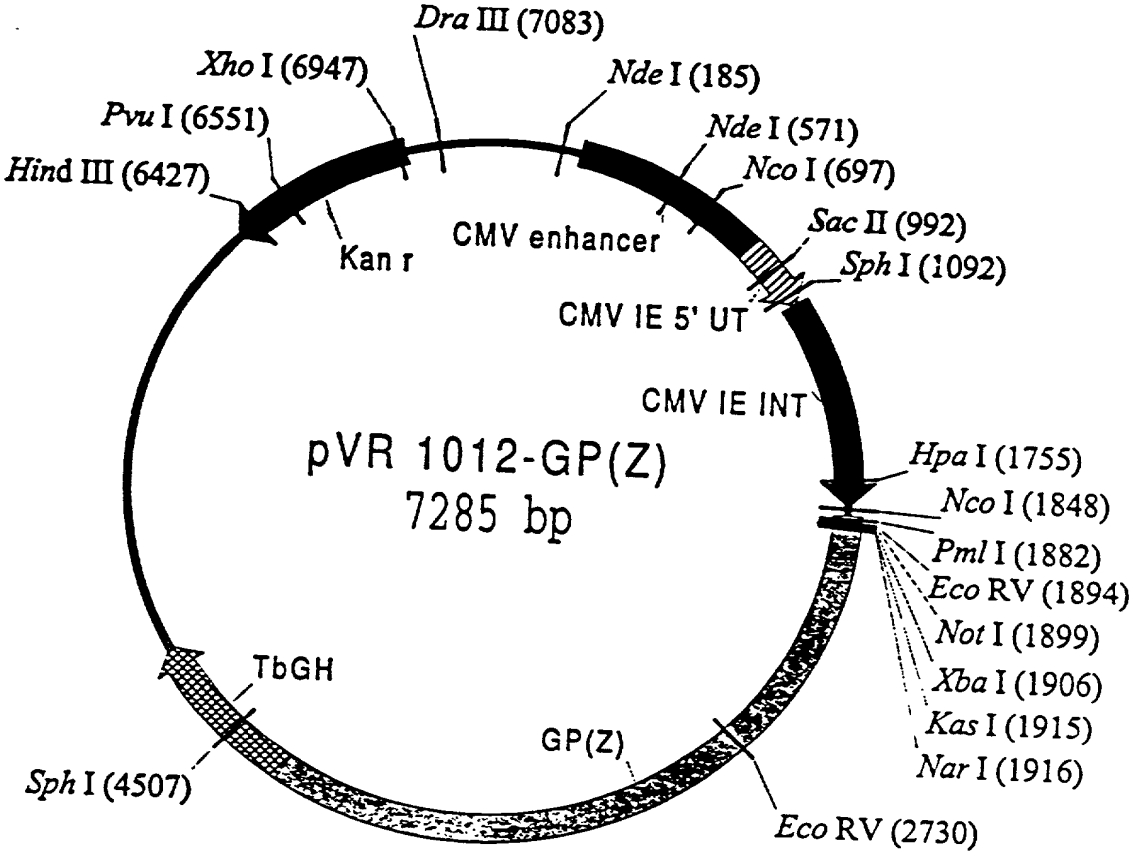


Figure 8

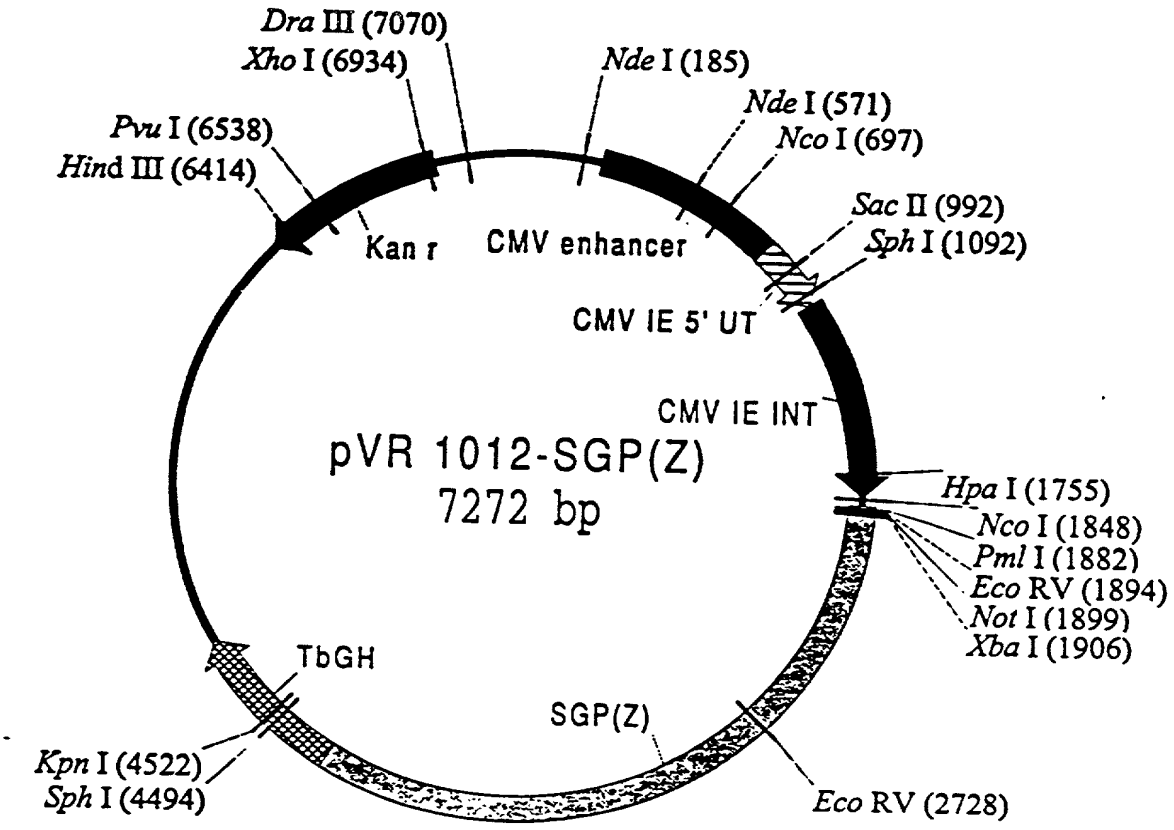

Dra III (7070)
Xho I (6934)
Nde I (185)
Nde I (571)
Nco I (697)
Pvu I (6538)
Hind III (6414)
Sac II (992)
Sph I (1092)
Kan r
CMV enhancer
CMV IE 5' UT
CMV IE INT
pVR 1012-SGP(Z)
7272 bp
Hpa I (1755)
Nco I (1848)
Pml I (1882)
Eco RV (1894)
Not I (1899)
Xba I (1906)
TbGH
SGP(Z)
Kpn I (4522)
Sph I (4494)
Eco RV (2728)

Figure 9

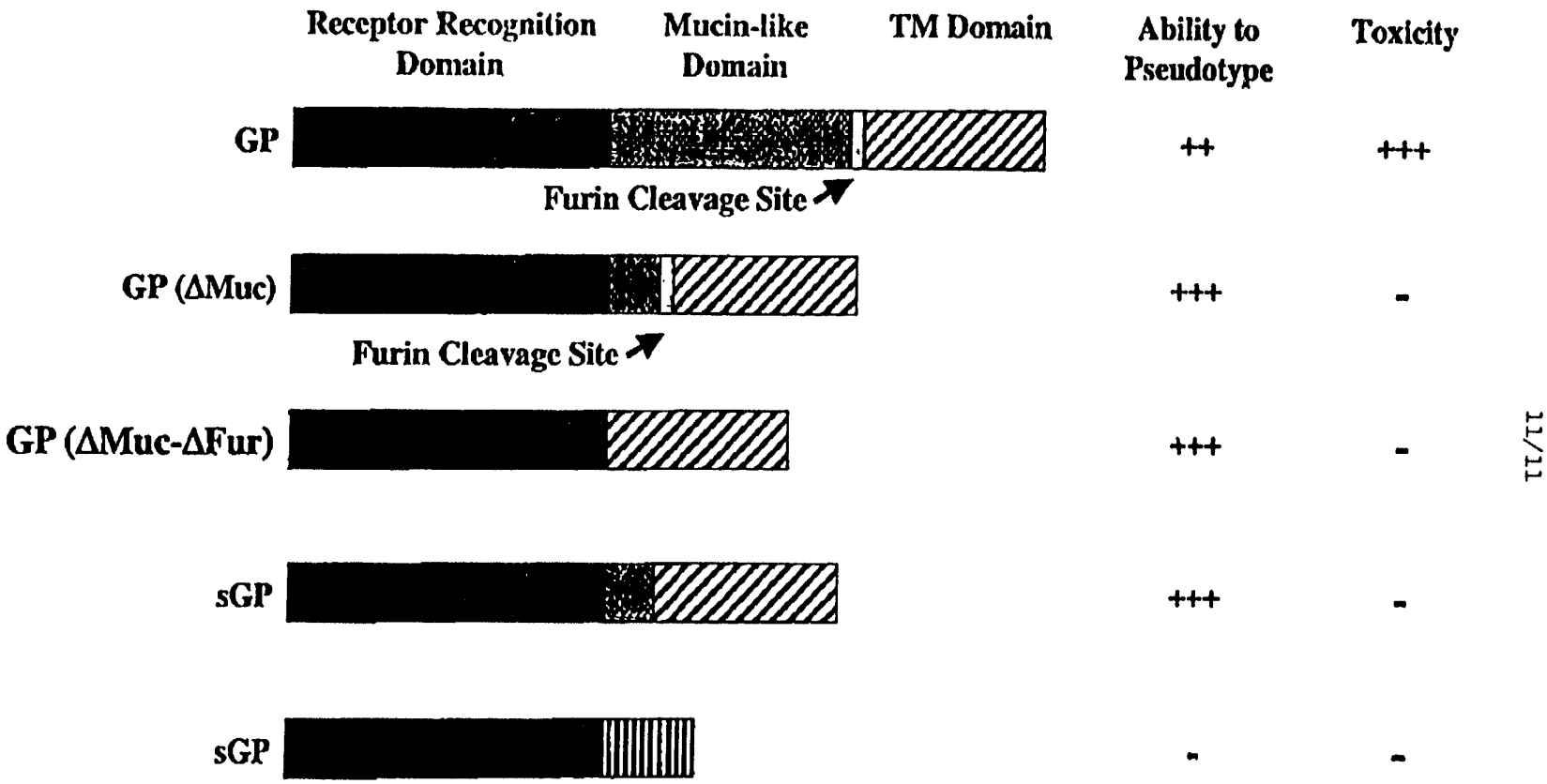


Figure 10

Atty Docket No.: **UMV-1474US**

DECLARATION, PETITION AND POWER OF ATTORNEY FOR PATENT APPLICATION

(Check one):

☐ Declaration Submitted with Initial Filing

☒ Declaration Submitted after Initial Filing

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

TARGETING GENE TRANSFER VECTORS TO CERTAIN CELL TYPES BY PSEUDOTYPING WITH VIRAL GLYCOPROTEIN

the specification of which (check one):

☐ is attached hereto.

OR

☒ was filed on **21 January 1999** as PCT International Application Number **PCT/US99/01382**

☐ and was amended by PCT Article 19 Amendment on _____ (if applicable),

☐ and was amended by PCT Article 34 Amendment on _____ (if applicable).

I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby state that I have reviewed and understood the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

PRIORITY CLAIM

(Check one):

☐ no such applications have been filed.

☒ such applications have been filed as follows

1) FOREIGN PRIORITY CLAIM: I hereby claim foreign priority benefits under Title 35, United States Code, §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate or §365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (dd/mm/yyyy)	Priority Not Claimed	Certified Copy Attached Yes	No
			☐	☐	☒
			☐	☐	☒

☐ Additional foreign application numbers are listed on a supplemental priority sheet attached hereto.

2) PROVISIONAL PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

Provisional Application Number(s)	Filing Date (dd/mm/yyyy)
60/072,033	**21 January 1998 (21.01.98)**

☐ Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

3) U.S./PCT PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code, §120 of any United States application or §365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose information which is known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application Number	PCT Parent Number	Parent Filing Date (dd/mm/yyyy)	Parent Patent Number *(if applicable)*
	PCT/US99/01382	**21 January 1999 (21.01.99)**	

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority

DECLARATION, PETITION AND POWER OF ATTORNEY FOR PATENT APPLICATION

(Check one):

☐ Declaration Submitted with Initial Filing

☒ Declaration Submitted after Initial Filing

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

TARGETING GENE TRANSFER VECTORS TO CERTAIN CELL TYPES BY PSEUDOTYPING WITH VIRAL GLYCOPROTEIN

the specification of which (check one):

☐ is attached hereto.

OR

☒ was filed on **21 January 1999** as PCT International Application Number **PCT/US99/01382**

☐ and was amended by PCT Article 19 Amendment on _____ (if applicable),

☐ and was amended by PCT Article 34 Amendment on _____ (if applicable).

I acknowledge the duty to disclose to the Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby state that I have reviewed and understood the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

PRIORITY CLAIM

(Check one):

☐ no such applications have been filed.

☒ such applications have been filed as follows

1) FOREIGN PRIORITY CLAIM: I hereby claim foreign priority benefits under Title 35, United States Code, §119(a)-(d) or §365(b) of any foreign application(s) for patent or inventor's certificate or §365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (dd/mm/yyyy)	Priority Not Claimed	Certified Copy Attached Yes	Certified Copy Attached No
			☐	☐	☒
			☐	☐	☒

☐ Additional foreign application numbers are listed on a supplemental priority sheet attached hereto.

2) PROVISIONAL PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code §119(e) of any United States provisional application(s) listed below.

Provisional Application Number(s)	Filing Date (dd/mm/yyyy)
60/072,033	**21 January 1998 (21.01.98)**

☐ Additional provisional application numbers are listed on a supplemental priority sheet attached hereto.

3) U.S./PCT PRIORITY CLAIM: I hereby claim the benefit under Title 35, United States Code, §120 of any United States application or §365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose information which is known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application Number	PCT Parent Number	Parent Filing Date (dd/mm/yyyy)	Parent Patent Number *(if applicable)*
	PCT/US99/01382	**21 January 1999 (21.01.99)**	

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority

POWER OF ATTORNEY:

As a named inventor, I hereby appoint the following attorneys and/or agents to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

James E. Cockfield	Reg. No. 19,162	Megan E. Williams	Reg. No. 43,270
Thomas V. Smurzynski	Reg. No. 24,798	Nicholas P. Triano III	Reg. No. 36,397
Ralph A. Loren	Reg. No. 29,325	Peter C. Lauro	Reg. No. 32,360
Giulio A. DeConti, Jr.	Reg. No. 31,503	DeAnn F. Smith	Reg. No. 36,683
Ann Lamport Hammitte	Reg. No. 34,858	William D. DeVaul	Reg. No. 42,483
Elizabeth A. Hanley	Reg. No. 33,505	David J. Rikkers	Reg. No. 43,882
Amy E. Mandragouras	Reg. No. 36,207	Chi Suk Kim	Reg. No. 42,728
Anthony A. Laurentano	Reg. No. 38,220	Maria Laccotripe Zacharakis	Limited Recognition Under 37 C.F.R. § 10.9(b)
Jane E. Remillard	Reg. No. 38,872		
Jeremiah Lynch	Reg. No. 17,425	Debra J. Milasincic	Reg. No. P46,931
Kevin J. Canning	Reg. No. 35,470	David R. Burns	Reg. No. P46,590
David A. Lane, Jr.	Reg. No. 39,261		
Jeanne M. DiGiorgio	Reg. No. 41,710		

Send Correspondence to:

DeAnn F. Smith, Lahive & Cockfield, LLP, 28 State Street, Boston, Massachusetts 02109, United States of America

Direct Telephone Calls to: (name and telephone number)

DeAnn F. Smith, (617) 227-7400

Wherefore I petition that letters patent be granted to me for the invention or discovery described and claimed in the attached specification and claims, and hereby subscribe my name to said specification and claims and to the foregoing declaration, power of attorney, and this petition.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor **Gary J. NABEL**
Inventor's signature Date
Residence **2520 30th Street, N.W., Washington, D.C. 20008**
Citizenship **United States of America**
Post Office Address (if different)

POWER OF ATTORNEY:

As a named inventor, I hereby appoint the following attorneys and/or agents to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

James E. Cockfield	Reg. No. 19,162	Megan E. Williams	Reg. No. 43,270
Thomas V. Smurzynski	Reg. No. 24,798	Nicholas P. Triano III	Reg. No. 36,397
Ralph A. Loren	Reg. No. 29,325	Peter C. Lauro	Reg. No. 32,360
Giulio A. DeConti, Jr.	Reg. No. 31,503	DeAnn F. Smith	Reg. No. 36,683
Ann Lamport Hammitte	Reg. No. 34,858	William D. DeVaul	Reg. No. 42,483
Elizabeth A. Hanley	Reg. No. 33,505	David J. Rikkers	Reg. No. 43,882
Amy E. Mandragouras	Reg. No. 36,207	Chi Suk Kim	Reg. No. 42,728
Anthony A. Laurentano	Reg. No. 38,220	Maria Laccotripe Zacharakis	Limited Recognition Under 37 C.F.R. § 10.9(b)
Jane E. Remillard	Reg. No. 38,872		
Jeremiah Lynch	Reg. No. 17,425	Debra J. Milasincic	Reg. No. P46,931
Kevin J. Canning	Reg. No. 35,470	David R. Burns	Reg. No. P46,590
David A. Lane, Jr.	Reg. No. 39,261		
Jeanne M. DiGiorgio	Reg. No. 41,710		

Send Correspondence to:

DeAnn F. Smith, Lahive & Cockfield, LLP, 28 State Street, Boston, Massachusetts 02109, United States of America

Direct Telephone Calls to: (name and telephone number)

DeAnn F. Smith, (617) 227-7400

Wherefore I petition that letters patent be granted to me for the invention or discovery described and claimed in the attached specification and claims, and hereby subscribe my name to said specification and claims and to the foregoing declaration, power of attorney, and this petition.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of sole or first inventor **Gary J. NABEL**
Inventor's signature Gary J Nabel Date 4/25/01
Residence **2520 30th Street, N.W., Washington, D.C. 20008** U.S.
Citizenship **United States of America**
Post Office Address (if different)

2W

Full name of sole or first inventor **Rafael DELGADO**	
Inventor's signature R Delgado	Date
Residence PASEO DE SOMOSIERRA 30 TRES CANTOS ; 28761 , MADRID , SPAIN ESX	
Citizenship **Spain**	
Post Office Address (if different)	

Full name of sole or first inventor **Zhi-yong YANG**	
Inventor's signature	Date
Residence **10004 Penfold Ct., Potomac, Maryland 20854**	
Citizenship **China**	
Post Office Address (if different)	

Full name of sole or first inventor **Rafael DELGADO**
Inventor's signature Date
Residence
Citizenship **Spain**
Post Office Address (if different)

Full name of sole or first inventor **Zhi-yong YANG**
Inventor's signature Date 3/26/2001
Residence **10004 Penfold Ct., Potomac, Maryland 20854** U.S.
Citizenship **China**
Post Office Address (if different)

3W